

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 17:55:45 ; Search time 29 Seconds (without alignments)

113.681 Million cell updates/sec

Title: US-10-006-223-1

Perfect score: 82

Sequence: VVGGTRAAQGEFPFMV 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

1 number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

-processing: Minimum Match 0% Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*

1: sp_archaea:*

2: sp_bacteria:*

3: sp_fungi:*

4: sp_human:*

5: sp_invertebrate:*

6: sp_mammal:*

7: sp_mhc:*

8: sp_organelle:*

9: sp_plant:*

10: sp_rabbit:*

11: sp_rat:*

12: sp_virus:*

13: sp_vertebrate:*

14: sp_unclassified:*

15: sp_rvirus:*

16: sp_bacteriop:*

17: sp_archeap:*

17 52 63.4 548 16 Q9KRU1
 18 51 62.2 262 5 Q9VRF9
 19 51 62.2 405 5 Q44331
 Q8wvc1 homo sapien
 Q8r055 mus musculus
 001310 bottlylus s.
 Q9vrs6 drosophila
 Q9vbs drosophila
 Q9v46 drosophila
 Q9v4n6 drosophila
 Q9xy6 rhypopertha
 Q28609 oryctolagus
 Q9vvt3 drosophila
 Q9v900 scolopendra
 Q9i7j3 drosophila
 Q93711 xenopus lae.
 Q9dgri xenopus lae.
 P81802 eisenia foetida
 Q9ven0 drosophila
 Q9vvt3 drosophila
 Q8t3a2 ciona intestinalis
 Q9xy58 tenocephalites
 Q9ty16 penaeus vannamei
 Q27761 penaeus vannamei
 Q96899 scolopendra
 Q9gqz4 rattus norvegicus
 P79343 bos taurus
 Q8vhj4 rattus norvegicus
 Q9a500 caulobacter
 Q9tr3 oryctolagus

ALIGNMENTS

RESULT 1
 Q54168 ID PRELIMINARY; PRT; 259 AA.
 AC Q54168; DT 01-NOV-1996 (TRMBLrel. 01, Created)
 DT 01-NOV-1996 (TRMBLrel. 01, Last sequence update)
 DT 01-MAR-2002 (TRMBLrel. 20, Last annotation update)
 DE Trypsinogen.
 OS Streptomyces fradiae.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetaceae; Streptomyces; Streptomyces.
 OC Actinomycetaceae; Streptomyces; Streptomyces; Streptomyces.
 OC NCBI_TaxID=1906;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 14544;
 RA Katoh T., Kikuchi N., Nagata K., Yoshida N.;
 RT "Cloning and expression of Trypsin-like enzyme from Streptomyces
 RT fradiae for comparative analysis of functional regions of Streptomyces
 RT and Mammalian Trypsins";
 RL J. Ferment. Bioeng. 80:440-445 (1995).
 CC - - SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPSIN FAMILY.
 DR EMBL; D16687; BAA04089.1; -.
 DR HSSP; P00775; 1SGT.
 DR MEROPS; S01.101; -.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Serine protease_Try.
 DR PFAM; PF00088; trypsin; 1.
 DR PRINTS; PRO0022; CHYNOTRYPSIN.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00240; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00134; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease.
 FT CHAIN 39 TRYPSIN.
 SQ SEQUENCE 259 AA; 26702 MW; 8B6C8496AB840FCF CRC64;
 Query Match 93.9%; Score 77; DB 2; Length 259;
 Best Local Similarity 93.8%; Pred. No. 1.8e-05;

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	77	93.9	259	2	Q54168	054168 streptomyces	
2	62	75.6	255	3	Q9y7a9	Q9y7a9 metarhizium	
3	62	75.6	261	3	Q00344	Q00344 cochliobolus	
4	61	74.4	311	3	Q8WZM5	Q8WZM5 trichoderma	
5	60	73.2	254	3	Q01136	Q01136 metarhizium	
6	60	73.2	256	3	Q9Y842	Q9Y842 metarhizium	
7	59	72.0	265	3	Q74696	Q74696 phaeosphaer	
8	56	68.3	360	5	Q9v1x6	Q9v1x6 drosophila	
9	55	67.1	281	5	Q76898	Q76898 drosophila	
10	54	65.9	242	5	Q96687	Q96687 lumbricus	
11	54	65.9	251	5	Q9vxc9	Q9vxc9 drosophila	
12	54	65.9	294	5	Q9vtr3	Q9vtr3 drosophila	
13	54	65.9	974	13	Q90WD8	Q90WD8 bufo japonicus	
14	53	64.6	268	5	Q46151	Q46151 pacifastacus	
15	53	64.6	855	11	Q9qj17	Q9qj17 rattus norvegicus	
16	52	63.4	155	5	Q95V74	Q95V74 homarus americanus	

RA	Mount S.M., Moy M., Murphy B., Nixon L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon G., Nusskern D.R., Paciob J.M., Pollard J., Puri V., Reese M.G., Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H., Spier E., Spradling A.C., Simpson M., Skupski M.P., Smith T., Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X., Wang Z.-Y., Wasserman D.A., Weinstock J., Weissbach J., Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao O.A., Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L., Gibbs R.A., Myers E.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O., RA	Query Match 65.9%; Score 54; DB 13; Length 974; Best Local Similarity 62.5%; Pred. No. 0.85; Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
RA	Qy 1 VGGTRAQGEPEPMV 16 Db 50 IVGGTSVAKGESPMV 65	Qy 1 VGGTRAQGEPEPMV 16 Db 50 IVGGTSVAKGESPMV 65
RA	RESULT 14	RESULT 14
DR	046151 PRELIMINARY; PRT; 268 AA.	046151 PRELIMINARY; PRT; 268 AA.
DR	046151, Created	046151, Created
DR	01-JUN-1998 (TREMBLrel. 21, Last annotation update)	01-JUN-1998 (TREMBLrel. 21, Last annotation update)
DR	01-JUN-2002 (TREMBLrel. 21, Last sequence update)	01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DR	DE Trypsin precursor (EC 3.4.21.4) (Fragment).	DE Trypsin precursor (EC 3.4.21.4) (Fragment).
DR	OS Pacificastacus leniusculus (Signal, crayfish).	OS Pacificastacus leniusculus (Signal, crayfish).
DR	RA Hernandez-Cortes M.P., Cerenius L., Garcia-Carreno F.L., Soderhall K., Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Bivalvia; Decapoda; Astacidae; Astacoidea; Astacidae; Pacifastacus.	RA Hernandez-Cortes M.P., Cerenius L., Garcia-Carreno F.L., Soderhall K., Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Bivalvia; Decapoda; Astacidae; Astacoidea; Pacifastacus.
DR	NCBI_TaxID=6720;	NCBI_TaxID=6720;
RN	[1] Sequence from N.A.	[1] Sequence from N.A.
RC	TISSUE=HEPATOPANCREAS;	TISSUE=HEPATOPANCREAS;
RX	MERLINE-99281768; Published=10355637;	MERLINE-99281768; Published=10355637;
RA	RA Hernandez-Cortes M.P., Cerenius L., Garcia-Carreno F.L., Soderhall K., Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Bivalvia; Decapoda; Astacidae; Astacoidea; Pacifastacus.	RA Hernandez-Cortes M.P., Cerenius L., Garcia-Carreno F.L., Soderhall K., Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca; Eumalacostraca; Bivalvia; Decapoda; Astacidae; Astacoidea; Pacifastacus.
DR	RT "Trypsin from Pacificastacus leniusculus Hepatopancreas: Purification and cDNA Cloning of the Synthesized Zymogen."	RT "Trypsin from Pacificastacus leniusculus Hepatopancreas: Purification and cDNA Cloning of the Synthesized Zymogen."
RL	RL Biol. Chem. 380:499-501 (1999).	RL Biol. Chem. 380:499-501 (1999).
CC	CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY.	CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE TRYPsin FAMILY.
DR	DR EMBL; AJ222658; CAA10915.1; -.	DR EMBL; AJ222658; CAA10915.1; -.
DR	DR HSSP; P00765; 1DPO.	DR HSSP; P00765; 1DPO.
DR	DR MEROPS; S01.122; -.	DR MEROPS; S01.122; -.
DR	DR InterPro; IPR001969; Asprotease site.	DR InterPro; IPR001969; Asprotease site.
DR	DR InterPro; IPR001314; Chymotrypsin.	DR InterPro; IPR001314; Chymotrypsin.
DR	DR InterPro; IPR001254; Ser_protease_Try.	DR InterPro; IPR001254; Ser_protease_Try.
DR	DR Pfam; PF00089; trypsin; 1.	DR Pfam; PF00089; trypsin; 1.
DR	DR PRINTS; PR00722; CHYMOTRYPSIN.	DR PRINTS; PR00722; CHYMOTRYPSIN.
DR	DR SMART; SM00000; TRYSPN_SPC; 1.	DR SMART; SM00000; TRYSPN_SPC; 1.
DR	DR PROSITE; PS00020; TRYSPN_SPEC; 1.	DR PROSITE; PS00020; TRYSPN_SPEC; 1.
DR	DR PROSITE; PS50240; TRYSPN_DOM; 1.	DR PROSITE; PS50240; TRYSPN_DOM; 1.
DR	DR PROSITE; PS00134; TRYSPN_HIS; 1.	DR PROSITE; PS00134; TRYSPN_HIS; 1.
DR	DR PROSITE; PS00135; TRYSPN_SER; 1.	DR PROSITE; PS00135; TRYSPN_SER; 1.
KW	KW Hydrolase; Serine protease.	KW Hydrolase; Serine protease.
SQ	SEQUENCE 294 AA; 32482 MW; 4867AGC355BB3C51 CRC64;	SEQUENCE 294 AA; 32482 MW; 4867AGC355BB3C51 CRC64;
Query Match 65.9%; Score 54; DB 5; Length 294; Best Local Similarity 66.7%; Pred. No. 0.23; Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	Query Match 65.9%; Score 54; DB 5; Length 294; Best Local Similarity 66.7%; Pred. No. 0.23; Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;	
DR	DR [1] Sequence from N.A.	DR [1] Sequence from N.A.
DR	DR TISSUE=OVIDUCTAL PARS RECTA;	DR TISSUE=OVIDUCTAL PARS RECTA;
RA	RA Hiyoshi M., Takamure K., Mita K., Kubo H., Sugimoto Y., Katagiri C., "Oriductin, the oviductal protease that mediates gamete interaction by affecting the vitelline envelope in <i>Bufo japonicus</i> : Its molecular cloning and analyses of expression and post-translational activation.", Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.	RA Hiyoshi M., Takamure K., Mita K., Kubo H., Sugimoto Y., Katagiri C., "Oriductin, the oviductal protease that mediates gamete interaction by affecting the vitelline envelope in <i>Bufo japonicus</i> : Its molecular cloning and analyses of expression and post-translational activation.", Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RL	RL MEROPS; S01.240; -.	RL MEROPS; S01.240; -.
DR	DR InterPro; IPR00855; CUB_domain.	DR InterPro; IPR00855; CUB_domain.
DR	DR Pfam; PF00431; CUB; 2.	DR Pfam; PF00431; CUB; 2.
DR	DR Pfam; PF0089; trypsin; 2.	DR Pfam; PF0089; trypsin; 2.
DR	DR PROSITE; PS001180; TRYSPN_DOM; 2.	DR PROSITE; PS001180; TRYSPN_DOM; 2.
DR	DR PROSITE; PS50240; TRYSPN_DOM; 2.	DR PROSITE; PS50240; TRYSPN_DOM; 2.
DR	DR PROSITE; PS00134; TRYSPN_HIS; UNKNOWN_1.	DR PROSITE; PS00134; TRYSPN_HIS; UNKNOWN_1.
KW	KW Hydrolase; Serine protease.	KW Hydrolase; Serine protease.
SQ	SEQUENCE 974 AA; 107647 MW; F19705RA470465553 CRC64;	SEQUENCE 974 AA; 107647 MW; F19705RA470465553 CRC64;
Query Match 64.6%; Score 53; DB 5; Length 268; Best Local Similarity 56.2%; Pred. No. 0.32; Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;	Query Match 64.6%; Score 53; DB 5; Length 268; Best Local Similarity 56.2%; Pred. No. 0.32; Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;	
DR	DR [1] VUGGTRAQGEPEPYQL 47	DR [1] VUGGTRAQGEPEPYQL 47
DR	DR Q9JUJ7 PRELIMINARY; PRT; 855 AA.	DR Q9JUJ7 PRELIMINARY; PRT; 855 AA.
DR	DR Q9JUJ7, Created	DR Q9JUJ7, Created
DR	DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)	DR 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DR	DR 01-JUN-2002 (TREMBLrel. 21, Last annotation update)	DR 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DR	DR Membrane bound serine protease (Membrane bound arginine specific serine protease).	DR Membrane bound serine protease (Membrane bound arginine specific serine protease).
GN	GN Rattus norvegicus (Rat)	GN Rattus norvegicus (Rat)
OC	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Butheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=JEJUNUM;
 RA Tsuzuki S.;
 RR "A membrane bound serine protease expressed in rat small intestine."
 RL Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=WISTAR; TISSUE=DUODENUM;
 RA Inoue H., Takahashi K., Kishi K.;
 RP membrane-bound arginine specific serine protease."
 RL Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
 CC TRYPsin FAMILY.
 DR EMBL: AB037698; BAB03502.1; -.
 DR EMBL: AB049189; BAB13765.1; -.
 DR HSMP; P00763; 1DPO.
 DR MEROPS: S01.302; -.
 DR InterPro; IPR002106; AACRNA_ligaseII.
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR000559; CUB domain.
 DR InterPro; IPR002172; LDL_recept_A.
 DR InterPro; IPR001354; Ser_protease_Try.
 PFam; PF00431; CUB; 2.
 PFam; PF00057; LDL_recept_a; 4.
 DR PRINTS; PR00722; CHYMOTRYPSIN.
 DR PRINTS; PR00261; LDLRECEPTOR.
 DR SMART; SM00042; CUB; 2.
 DR SMART; SM00192; LDLa; 3.
 DR SMART; SM00020; TRYPSIN; 1.
 DR PROSITE; PS00339; AA_TRNA_DIGASE_II_2; UNKNOWN_1.
 DR PROSITE; PS001180; CUB; 2.
 DR PROSITE; PS001209; LDLra_1; 2.
 DR PROSITE; PS50068; LDLra_2; 4.
 DR PROSITE; PS50240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00134; TRYPSIN_HIS; UNKNOWN_1.
 DR PROSITE; PS00135; TRYPSIN_SBR; 1.
 KW Glycoprotein; Hydrolase; Protease; Serine protease.
 FT VARIANT 665 665 K->N.
 SQ SEQUENCE 855 AA; 94955 MW; 3580687ECF6CF03D CRC64;
 Query Match 64.6%; Score 53; DB 11; Length 855;
 Best Local Similarity 62.5%; Pred. No. 1.1;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 1 VGGTRAAQGEPPFMV 16
 615 VGGTNADEGENPWQV 630

Search completed: April 11, 2003, 17:58:07
 time : 31 secs



Gendore version 5.1.4 p5_4578 (c) 1993 - 2003 Compugen Ltd.											
protein - protein search, using sw model											
run on:	April 11, 2003, 17:55:20 ;	Search time 11 Seconds (without alignments)									
		60.329 Million cell updates/sec									
title:	US-10-006-223-1										
perfect score:	82										
sequence:	1 VVGTRAAQGEFPFPMV 16										
conning table:	BLOSUM62										
	Gapop 10.0 , Gapext 0.5										
searched:	112892 seqs, 41476328 residues										
	112892 seqs, 41476328 residues										
number of hits satisfying chosen parameters:	112892										
	112892 seqs, 41476328 residues										
-processing:	Minimum Match 0%										
	Maximum Match 100%										
	Listing first 45 summaries										
	SwissProt_40:*										
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.											
SUMMARIES											
		Query	Match	Length	ID	Description					
result No.	Score										
1	82	100.0	259	1	TRYP_STRGR	P00775 streptomyces halotilis					
2	62	75.6	268	1	TRYP_STRGR	Q54179 fusarium					
3	59	72.0	248	1	TRYP_FUSOX	P35049 streptomyces					
4	53	70.7	20	1	TRYL_STREX	P80420 streptomyces					
5	53	64.6	855	1	STYL4_MOUSE	P56677 mus musculus					
6	52	63.4	216	1	CR2_YESSOR	P00768 vespa orientalis					
7	52	63.4	237	1	TRYP_ASTFL	P00765 astacus fluviatilis					
8	52	63.4	254	1	CPRL_HALRU	P35003 ovis aries					
9	52	63.4	274	1	FA9_SHEEP	P16293 oryctolagus cuniculus					
10	51	62.2	638	1	KAL_MOUSE	P26762 mus musculus					
11	51	62.2	855	1	STYL7_HUMAN	Q9Y9Y6 homo sapiens					
12	50	61.0	32	1	TRYP_PENMO	P35050 penaeus japonicus					
13	50	61.0	218	1	CTR2_VESCR	P00769 vespa crabro					
14	49	59.8	219	1	CPBP7_FIG	P35001 subterraneus					
15	49	59.8	431	1	ACRO_RABIT	P48038 oryctolagus cuniculus					
16	49	59.8	625	1	FP11_BOVIN	P03951 homo sapiens					
17	48	58.5	20	1	CGGC_PARCM	P20134 paralithoeces					
18	48	58.5	271	1	S24D_ANOGA	P17604 anophelis gambiae					
19	48	58.5	638	1	KAL_HUMAN	P03952 homo sapiens					
20	48	58.5	638	1	KAL_RAT	P14222 rattus norvegicus					
21	47	57.3	60	1	ACRO_CAPIH	P10626 capra hircus					
22	47	57.3	416	1	FA9_BOVIN	P00741 bos taurus					
23	46	56.1	271	1	FA9_PIG	P16293 sus scrofa					
24	46	56.1	275	1	FA9_RABIT	P16292 oryctolagus cuniculus					
25	46	56.1	705	1	C1R_HUMAN	P00736 homo sapiens					
26	45	54.9	20	1	CGC2_CHIOP	P34154 chionoecetes					
27	45	54.9	20	1	CGG3_CHIOP	P34155 chionoecetes					
28	45	54.9	20	1	CGGA_PARCM	P20732 paralithoeces					
29	45	54.9	20	1	CGCB_PARCM	P20733 paralithoeces					
30	45	54.9	267	1	ELNE_HUMAN	P00246 homo sapiens					
31	45	54.9	285	1	FA9_CAVPO	P16295 cavia porcellus					
32	45	54.9	418	1	HATI_HUMAN	060235 homo sapiens					
33	45	54.9	422	1	DESI_HUMAN	091152 homo sapiens					

34	45	54.9	452	1	FA9 CANFA	P19540 canis famili
		54.9	461	1	FA9_HUMAN	P00740 homo sapien
		53.7	261	1	DBR3_DERPT	P3675 dermatophag
		53.7	261	1	EUM3_EURM	Q97370 europlymph
		53.7	583	1	CFA1_HUMAN	P05156 homo sapien
		53.7	1019	1	ENTK_HUMAN	P98073 homo sapien
		53.7	251	1	CAP7_HUMAN	P02160 homo sapien
		52.4	415	1	ACRO_PIG	P08001 sus scrofa
		52.4	421	1	ACRO_HUMAN	P101323 homo sapien
		52.4	490	1	TMS2_MOUSE	Q9jgq8 mus musculi
		52.4	604	1	CFA1_RAT	Q9w398 rattus norv
		52.4	686	1	MAS2_HUMAN	Q99118 macaca mulatta

ALIGNMENTS

RESULT 1					
	TRYP	STRGR	STANDARD	PRT;	259 AA.
ID	TRYP	STRGR			
AC	P00775;				
DT	21-JUL-1986	(Rel. 01, Created)			
DT	01-FEB-1994	(Rel. 28, Last sequence update)			
DT	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Trypsin precursor (EC 3.4.21.4) (SGT).				
GN	SPTK.				
OS	Streptomyces griseus.				
OC	Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;				
OC	Actinomycetales; Streptomyceae; Streptomyces.				
OX	NCBIXN				
TX	TaxId=91;				

[1] SEQUENCE FROM N.A.
STRAIN=ANCC 10137;
MEDLINE=92095577; PubMed=1755852;
Kim J.C., Cha S.H., Jeong S.T., Oh S.K., Byun S.M.;
"Molecular cloning and nucleotide sequence of *Streptomyces griseus*
trypsin gene.",
Biochem. Biophys. Res. Commun. 181:707-713 (1991).
[2] SEQUENCE OF 37-259.
MEDLINE=75127940; PubMed=804314;
Olafson R.W., Jurasek L., Carpenter M.R., Smillie L.B.;
"Amino acid sequence of *Streptomyces griseus* trypsin. Cyanogen
bromide fragments and complete sequence.",
Biochemistry 14:1168-1177 (1975).
[3] X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
MEDLINE=88286735; PubMed=3135412;
Read R.J., James M.N.G.;
"Refined crystal structure of *Streptomyces griseus* trypsin at 1.7-A
resolution.",
J. Mol. Biol. 200:523-551 (1988).
-|- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-xaa, Lys-|-xaa.
-|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation
of the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a licence agreement (see <http://www.isb-bib.ch/announce>
or send an email to license@isb-bib.ch).

EMBL; M64471; AA026820.1; ALT_SEQ.
PIR; A00962; TRSMG.
PIR; JO1302; JO1302.
PDB; 1SGT; 16-JUL-88.
MEROPS; S01.101; -.
InterPro; IPR01314; Chymotrypsin.
InterPro; IPR01254; Ser_protease_Try.
Pfam; PF00089; trypsin_1.
PRINTS; PRO0722; CHYMATRYSIN.
SMART; SM00220; TRYSPIN.
TMSRF; TMSRF.

DR PROSITE; PS50240; TRYPSIN DOM; 1.
 DR PROSITE; PS00134; TRYPSIN HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_BER; 1.
 KW Hydrolase; Serine protease; Zymogen; Signal; 3D-structure.

FT SIGNAL 1 32 ACTIVATION PEPTIDE.
 FT PROPEP 33 36 TRYPSIN.
 FT CHAIN 37 259 CHARGE RELAY SYSTEM.
 FT ACT_SITE 73 73 CHARGE RELAY SYSTEM.
 FT ACT_SITE 118 118 CHARGE RELAY SYSTEM.
 FT ACT_SITE 208 208 CHARGE RELAY SYSTEM.
 FT DISULFID 58 74 DISULFID 177 192 DISULFID 204 233 DISULFID 202 202 REQUIRED FOR SPECIFICITY.
 FT SITE 95 96 MISSING (IN REF. 2).
 FT CONFLICT 38 38 STRAND 38 38 STRAND 65 66 TURN 39 39 STRAND 41 42 TURN 45 46 TURN 49 50 STRAND 51 54 TURN 55 57 STRAND 58 64 TURN 65 66 STRAND 67 70 HELIX 72 74 STRAND 79 80 STRAND 85 88 STRAND 92 92 TURN 93 94 TURN 96 97 TURN 99 108 TURN 110 111 TURN 120 124 STRAND 134 135 TURN 140 141 STRAND 145 150 TURN 157 156 STRAND 163 163 STRAND 165 172 HELIX 174 184 TURN 187 189 STRAND 190 193 TURN 196 198 STRAND 202 202 TURN 205 206 STRAND 208 209 TURN 211 215 STRAND 217 218 TURN 221 229 STRAND 236 237 TURN 240 244 HELIX 245 257 TURN 258 258 SEQUENCE 259 AA; 26776 MW; 050233APP1F64823 CRC64;

FT SIGNAL 100.0%; Score 82; DB 1; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2.9e-07; Indels 0; Gaps 0;

FT STRAND 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FT STRAND 16; VVGGRRAQGEFPFMV 16
 FT STRAND 52; VVGGRRAQGEFPFMV 52

RESULT 2
 TRYPSIN STRGA STANDARD; PRT; 268 AA.
 AC Q54177;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin-like protease precursor (EC 3.4.21.-).

Query Match 1 VVGGRRAQGEFPFMV 16
 Mismatches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Db 46 VIGGKPAQNEFPFMV 61

RESULT 3
 TRYPSIN FUSOX STANDARD; PRT; 248 AA.
 ID P3049; AC P3049;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Fusarium oxyphorum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 NCBI_TaxID=507;

RN [1]; RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 ID TRYPSIN STRGA STANDARD; PRT; 268 AA.
 AC P3049;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Fusarium oxyphorum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 NCBI_TaxID=507;

RN [1]; RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 ID TRYPSIN FUSOX STANDARD; PRT; 248 AA.
 AC P3049;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin precursor (EC 3.4.21.4).
 OS Fusarium oxyphorum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 NCBI_TaxID=507;

RN [1]; RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS).
 ID TRYPSIN STRGA STANDARD; PRT; 268 AA.
 AC P3049;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Trypsin-like protease precursor (EC 3.4.21.-).

oxysporum."; Protein Eng. 6:341-348(1993).
 [2] X-RAY CRYSTALLOGRAPHY (1.55 ANGSTROMS).
 RA Rypniewski W.R., Dambmann C., von der Osten C., Dauter M., Wilson K.S.; "Structure of inhibited trypsin from *Fusarium oxysporum* at 1.55 Å.";
 RL Acta Crystallogr. D 51:73-84(1995).
 RT CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-xaa, Lys-|-xaa.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SUBCELLULAR LOCATION: Secreted.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A.
 CC -|- FUNCTION: INVOLVED IN MYCELIUM DIFFERENTIATION.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S2A.
 DR MEROPS: S01_1001; -
 DR InterPro: IPR001254; Ser_Protease_TRY.
 DR PROSITE: PS00134; TRYPSIN_HIS; PARTIAL.
 DR PROSITE: PS00135; TRYPSIN_SER; PARTIAL.
 KW Hydrolase; Serine protease.
 FT NON_TER 20
 SQ SEQUENCE 20 AA: 2129 MW: 4568F56D0E7393AF CRC64;
 Query Match 70.7%; Score 58; DB 1; Length 20;
 Best Local Similarity 84.6%; Pred. No. 0.0037;
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 2 VGGTTRAQGEPPF 14
 Db 2 VGGTTRAQGNFPF 14

RESULT 5
 ST14 MOUSE STANDARD; PRT; 855 AA.
 ID ST14 MOUSE STANDARD; PRT; 855 AA.
 AC P56677;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Suppressor of tumoractivity 14 (EC 3.4.21.-) (Epithin).
 ST14 OR PRSS14.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBITaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C.B.17SCID; TISSUE=Thymus;
 RX MEDLINE=99z16440; PubMed=10199918;
 RA Kim M.G., Chen C., Lyu M.S., Cho E.G., Park D., Kozak C., Schwartz R.H.;
 RA "Cloning and chromosomal mapping of a gene isolated from thymic
 strional cells encoding a new mouse type II membrane serine protease,
 epithin, containing four LDL receptor modules and two CUB domains."
 RT Immunogenetics 49:420-428(1999).
 RL [2]
 RP REVISIONS TO 23; 321; 325; 343; 409-410 AND C-TERMINUS.
 RC STRAIN=C.B.17SCID; TISSUE=Thymus;
 RA Kim M.G., Chen C., Cho E.G., Park D., Schwartz R.H.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAUSBERG R.;
 RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -|- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
 CC -|- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN INTESTINE, KIDNEY, LUNG,
 CC AND THYMUS. NOT EXPRESSED IN SKELETAL MUSCLE, LIVER, HEART,
 CC TESTIS, AND BRAIN.
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 CC -|- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
 CC -|- SIMILARITY: CONTAINS 2 CUB DOMAINS.
 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a licence agreement (see <http://www.ebi-sib.ch/announce/>
 or send an email to license@ebi-sib.ch).

RESULT 4
 STREX STANDARD; PRT; 20 AA.
 AC P80420;
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Trypsin-like protease (EC 3.4.21.-) (Fragment).
 OS Streptomyces exfoliatus (Streptomyces hyrogenans).
 OC Bacteria; Actinobacteria; Streptomyces (class); Actinobacteridae;
 OC Actinomycetes; Streptomycineae; Streptomyctaceae; Streptomyces.
 OX NCBI_TaxID=1905;
 RN [1]
 RP SEQUENCE.
 RC STRAIN=SMF13;
 RX MEDLINE=95221424; PubMed=7733739;
 RA Kim I.S., Lee K.J.;
 RT "Physiological roles of leupeptin and extracellular proteases in
 mycelium development of *Streptomyces exfoliatus* SMF13.";
 RL Microbiology 141:1017-1025(1995).

FT ACT SITE 45 45 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 96 96 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT ACT SITE 189 189 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 30 46 REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 FT DISULFID 159 174
 FT DISULFID 185 213
 FT SITE 183 25021 MW; REQUIRED FOR SPECIFICITY (BY SIMILARITY).
 SQ SEQUENCE 237 AA; 25021 MW; 407213B55022C76 CRC64;

Query Match 63.4%; Score 52; DB 1; Length 23; ;
 Best Local Similarity 56.2%; Pred. No. 0.057; ;
 Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0; Gaps 0;

Qy 1 VVGTRAAQGEPPFMV 16
 :||||| :||||| :
 Db 1 IVGGTDAVGEPPQL 16

RESULT 8
 CTRL_HALRU STANDARD; PRT; 254 AA.
 AC P3073 ;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC Chymotrypsin-like serine proteinase precursor (EC 3.4.21.-).
 CC Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Haliozoa; Metazoa; Mollusca; Gastropoda; Archaeogastropoda;
 CC Bivalvia; Caprinea; Ovis;
 CC NCBI_TaxID=6454;
 RN [1]
 RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
 RX TISSUE=Intestine; MEDLINE=93243624; PubMed=842947;
 RA Groppe J.C.; Morse D.E.;
 RT "Molluscan chymotrypsin-like protease: structure, localization, and
 substrate specificity";
 RL Arch. Biochem. Biophys. 305:159-169 (1993).
 CC 1- FUNCTION: SPECIFICITY SIMILAR TO CHYMOTRYPSIN.
 CC 1- ENZYME REGULATION: ACTIVATED BY AN AUTOCATALYTIC MECHANISM.
 CC 1- SUBUNIT: MONOMER.
 CC 1- SUBCELLULAR LOCATION: Extracellular.
 CC 1- TISSUE SPECIFICITY: EXPRESSED SPECIFICALLY IN THE DISTAL QUARTER
 CC OF THE INTESTINE.
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL: X71438; CA050572.1; -
 PIR; S35585; S35585.
 PIR; S32750; S32750.
 HSSP; P00763; 1DPO.
 DR MEROPS; S01_121; -
 DR InterPro; IPR001314; Chymotrypsin.
 DR InterPro; IPR001254; Ser_protease_Try.
 DR Pfam; PF00089; trypsin; 1.
 DR PRINTS; PRO0722; CHYMOTRYPSIN.
 SMART; SM00020.
 DR PROSITE; PS00240; TRYPSIN_DOM; 1.
 DR PROSITE; PS00133; TRYPSIN_HIS; 1.
 DR PROSITE; PS00135; TRYPSIN_SER; 1.
 KW Hydrolase; Serine protease; Glycoprotein; Signal; Zymogen.
 PT PROPEP 1 18 POTENTIAL.
 PT CHAIN 19 23 ACTIVATION PEPTIDE (POTENTIAL).
 PT ACT_SITE 24 254 CHYMOTRYPSIN-LIKE SERINE PROTEINASE.
 PT ACT_SITE 68 68 CHARGE RELAY SYSTEM (BY SIMILARITY).
 PT ACT_SITE 117 117 CHARGE RELAY SYSTEM (BY SIMILARITY).

FT ACT SITE 212 212 CHARGE RELAY SYSTEM (BY SIMILARITY).
 FT DISULFID 53 69 BY SIMILARITY.
 FT DISULFID 146 218 BY SIMILARITY.
 FT DISULFID 181 199 BY SIMILARITY.
 FT DISULFID 208 233 BY SIMILARITY.
 SQ SEQUENCE 254 AA; 27250 MW; ADAAA8A22BEFCBC CRC64;

Query Match 63.4%; Score 52; DB 1; Length 254;
 Best Local Similarity 64.3%; Pred. No. 0.062; ;
 Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEPPFMV 14
 :||||| :||||| :
 Db 24 IVGGTDAVGEPPW 37

RESULT 9
 FA9_SHEEP STANDARD; PRT; 274 AA.
 ID FA9_SHEEP
 AC P16291;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DB Coagulation factor IX (EC 3.4.21.22) (Christmas factor) (Fragment).
 GN F9.
 OS Ovis aries (Sheep).
 OC Bivalvia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 CC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90152675; PubMed=2303254;
 RA Sarkar G.; Koebel D.D.; Sommer S.S.;
 RT "Direct sequencing of the activation peptide and the catalytic domain
 of the factor IX gene in six species.";
 RL Genomics 6:133-143 (1990).
 CC 1- FUNCTION: FACTOR IX IS A VITAMIN K-DEPENDENT PLASMA PROTEIN THAT
 PARTICIPATES IN THE INTRINSIC PATHWAY OF BLOOD COAGULATION BY
 CONVERTING FACTOR X TO ITS ACTIVE FORM IN THE PRESENCE OF CA++
 CC IONS, PHOSPHOLIPIDS, AND FACTOR VIIA.
 CC 1- CATALYTIC ACTIVITY: Hydrolyzes one Arg-
 CC Ile bond in factor X.
 CC 1- SUBUNIT: FACTOR IX IS ACTIVATED BY FACTOR XA, WHICH EXCISES THE
 CC ACTIVATION PEPTIDE AND THUS PRODUCES A MOLECULE CONSISTING OF 2
 CC CHAINS HELD TOGETHER BY ONE OR MORE DISULFIDE BONDS.
 CC 1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC DR M26233; AAA31520.1; -
 CC DR HSSP; P16293; 1PFX.
 CC DR MEROPS; S01_214; -
 CC DR InterPro; IPR001254; Ser_protease_Try.
 CC DR InterPro; IPR000294; VitK_dep_GLA.
 CC DR Pfam; PF00089; trypsin; 1.
 CC DR SMART; SM00020; Tryp_SPC; 1.
 CC DR PROSITE; PS00011; TRYPSIN_DOM; 1.
 CC DR PROSITE; PS00134; TRYPSIN_HIS; 1.
 CC DR PROSITE; PS00135; TRYPSIN_SER; 1.
 CC KW Blood coagulation; Plasma; Serine protease; Calcium-binding;
 CC KW Hydrolase; Glycoprotein.
 CC DR PT NON_TBR 1 1
 CC DR ACT_SITE 89 89 CHARGE RELAY SYSTEM.
 CC DR ACT_SITE 137 137 CHARGE RELAY SYSTEM.
 CC DR ACT_SITE 233 233 CHARGE RELAY SYSTEM.

FT	CARBOHYD	25	25	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PFam; PF00089; trypsin; 1.
FT	CARBOHYD	35	35	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PRINTS; PR00005; APPLEDOMAIN.
FT	CARBOHYD	40	40	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	PRINTS; PR00722; CHMOTRYSIN.
FT	CARBOHYD	128	128	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	SMART; SM00233; APPLE; 4.
FT	NON_TER	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).	DR	SMART; SM0020; TRYPSIN; 1.
SQ	SEQUENCE	274 AA:	30595 MW;	D3617PC3B1D33E9B CRC64;	DR	PROSITE; PS00495; APPLE; 4.
Query Match	63.4%	Score 52;	DB 1;	Length 274;	DR	PROSITE; PS50240; TRYPSIN; DOM; 1.
Best Local Similarity	62.5%	Prod. No.	0.067;		DR	PROSITE; PS00134; TRYPSIN HIS; 1.
Matches 10;	Conservative	3;	Mismatches	0;	KW	Hydrolase; Serine protease; Glycoprotein; Plasma; Zymogen; Signal; Fibrinolysis; Blood coagulation; Inflammatory response; Liver;
Sequence	1	VVGCTRAAQEFPFMY 16	49	VVGCTDARGQEPWQV 64	KW	Repeat.
Qy					FT	SIGNAL 1 19
					FT	CHAIN 1 390
					FT	CHAIN 20 638
RESULT 10	KAL_MOUSE		STANDARD;	PRT;	PLASMA KALLIKREIN HEAVY CHAIN.	
ID	KAL_MOUSE				PLASMA KALLIKREIN LIGHT CHAIN.	
15	P2262					
ID	KAL_MOUSE					
01	01-MAY-1992	(Ref. 22, Created)				
	01-MAY-1992	(Ref. 22, Last sequence update)				
DT	01-JUN-2002	(Ref. 41, Last annotation update)				
DE	Plasma kallikrein precursor (EC 3.4.21.34) (Plasma prekallikrein)					
DE	(Kininogenin) (Fleicher factor).					
GN	KLK81 OR KLK3 OR PK.					
OS	Mus musculus (Mouse).					
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX	NCBI_TaxID=10090;					
RN	[1]					
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.					
RC	STAINA-BALB/-; TISSUE=Liver;					
RX	Medline=91090844; PubMed=224928;					
RA	Seidah N.G., Sawyer N., Hamelein J., Mion P., Beaubien G.,					
RA	Brachspap L., Rochement J., Mbikay M., Chretien M.;					
RT	"Mouse plasma kallikrein: cDNA structure, enzyme characterization, and comparison of protein and mRNA levels among species.";					
RL	DNA Cell Biol. 9:737-748 (1990).					
CC	-1- FUNCTION: THE ENZYME CLEAVES LYS-ARG AND ARG-SER BONDS. IT ACTIVATES, IN A RECIPROCAL REACTION, FACTOR XII AFTER ITS BINDING TO A NEGATIVELY CHARGED SURFACE. IT ALSO RELEASES BRAODYKININ FROM HMW KINININ AND MAY ALSO PLAY A ROLE IN THE RENIN-ANGIOTENSIN SYSTEM BY CONVERTING PRORENIN INTO RENIN.					
CC	-1- CATALYTIC ACTIVITY: Cleaves selectively Arg- -Xaa and Lys- -Xaa bonds, including Lys- -Arg and Arg- -Ser bonds in (human) kininogen to release bradykinin.					
CC	-1- SUBUNIT: THE ZYMOGEN IS ACTIVATED BY FACTOR XII, WHICH CLEAVES THE MOLECULE INTO A LIGHT CHAIN, WHICH ASSOCIATES WITH HMW KINININ. THESE CHAINS ARE LINKED BY ONE OR MORE DISULFIDE BONDS.					
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1. PLASMA KALLIKREIN SUBUNIT.					
CC	-1- SIMILARITY: CONTAINS 4 APPLE DOMAINS.					
CC	-----					
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).					
CC	-----					
CC	EMBL; M58588; AAA63393.1; -				RESULT 11	
DR	PIR; A36575; KOMSPL.				ST14_HUMAN STANDARD; PRT; 855 AA.	
DR	HSSP; P00750; IRTF.				ID Q95Y6; Q9HS350; Q9BS01; Q9HB36;	
DR	MEROPS; S01.21.2; -				AC 16-OCT-2001 (Rel. 40, Created)	
DR	MGD; MGI:102849; Klk3.				DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DR	InterPro; IPR00177; Apple.				DT 15-JUN-2002 (Rel. 41, Last annotation update)	
DR	InterPro; IPR001314; Chymotrypsin.				DE Suppressor of tumorigenicity 14 (EC 3.4.21. -) (Matriptase) (Membrane-type serine protease 1) (MSP1) (Prostasin) (Serine protease TAGD-15) (Tumor associated protease 1) (Tumor associated differentially-expressed gene-15 protein).	
DR	InterPro; IPR001254; PAN.				GN ST14 OR PRSS14 OR SNC19 OR TAGD15.	
DR	Pfam; PF00024; PAN; 4.				OS Homo sapiens (Human).	
DR	-----				OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Homo.	
DR	-----				OC Primates; Catarrhini; Hominidae; Homo.	
DR	NCBI_TaxID=9606; [1]				NCBI_TaxID=9606;	

SEQUENCE FROM N.A.
MEDLINE=9930581; PubMed=10373424;
RA Lin C.Y.; Anders J.; Johnson M.; Sang Q.A.; Dickson R.B.; Sang Q.A.; Dickson R.B.; a matrix-degrading serine
RT protease with trypsin-like activity.;
RL J. Biol. Chem. 274:18231-18236(1999);
RN [2]

SEQUENCE FROM N.A.
MEDLINE=99432178; PubMed=10500122;
RA Takeuchi T.; Shuman M.A.; Craik C.S.;
RT "Reverse biochemistry: Use of macromolecular protease inhibitors to
dissect complex biological processes and identify a membrane-type
serine protease in epithelial cancer and normal tissue.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:11034-11061(1999);
RN [3]

SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Tanimoto H.; Underwood L.J.; Wang Y.; Shigemasa K.; Parmley T.H.;
RT "Molecular cloning of a novel transmembrane serine protease expressed
in complex biological processes and identify a membrane-type
serine protease in epithelial cancer and normal tissue.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
RN [4]

SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Tanimoto H.; Underwood L.J.; Wang Y.; Shigemasa K.; Parmley T.H.;
RT Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN [5]

SEQUENCE OF 327-855 FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DDBJ databases.
RN [6]

SEQUENCE OF 340-564 FROM N.A.
RA Cao J.; Fan W.; Zheng S.;
RT "Genomic analysis of a novel human serine protease SNC19.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN [7]

CHARACTERIZATION
RC TISSUE=Milk;
RY MEDLINE=993582; PubMed=10373425;
RA Lin C.Y.; Anders J.; Johnson M.; Dickson R.B.;
RT "Purification and characterization of a complex containing matrix metalloprotease
and a Kunitz-type serine protease inhibitor from human milk.";
RL J. Biol. Chem. 274:18237-18242(1999);
CC -I- FUNCTION: DEGRADES EXTRACELLULAR MATRIX. PROPOSED TO PLAY A ROLE
IN BREAST CANCER INVASION AND METASTASIS. EXHIBITS TRYPSIN-LIKE
ACTIVITY AS DEFINED BY CLEAVAGE OF SYNTHETIC SUBSTRATES WITH ARG
OR LYS AS THE P1 SITE.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 CUB DOMAINS.

EMBL; AF118224; AAD42765.;
EMBL; AF133086; AAF00109.;
EMBL; AB030036; BAB20376.;
DR EMBL; AF057145; AAG15395.;
DR EMBL; BC005826; AA05826.;
DR HSSP; P00763; IDPO.
DR GenBank; HGNC; HGNC11344; ST14.
DR MIM; 606797.;
DR MEROPS; S01; 302.;
DR InterPro; IPR000859; CUB-domain.
DR InterPro; IPR001314; Chymotrypsin.

SEQUENCE FROM N.A.
MEDLINE=9930581; PubMed=10373424;
RA Lin C.Y.; Anders J.; Johnson M.; Sang Q.A.; Dickson R.B.; Sang Q.A.; Dickson R.B.; a matrix-degrading serine
RT protease with trypsin-like activity.;
RL J. Biol. Chem. 274:18231-18236(1999);
RN [2]

SEQUENCE FROM N.A.
MEDLINE=99432178; PubMed=10500122;
RA Takeuchi T.; Shuman M.A.; Craik C.S.;
RT "Reverse biochemistry: Use of macromolecular protease inhibitors to
dissect complex biological processes and identify a membrane-type
serine protease in epithelial cancer and normal tissue.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:11034-11061(1999);
RN [3]

SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Tanimoto H.; Underwood L.J.; Wang Y.; Shigemasa K.; Parmley T.H.;
RT "Molecular cloning of a novel transmembrane serine protease expressed
in complex biological processes and identify a membrane-type
serine protease in epithelial cancer and normal tissue.";
RL Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]

SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Tanimoto H.; Underwood L.J.; Wang Y.; Shigemasa K.; Parmley T.H.;
RT Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN [5]

SEQUENCE OF 327-855 FROM N.A.
RC TISSUE=Muscle;
RA Tanimoto H.; Underwood L.J.; Wang Y.; Shigemasa K.; Parmley T.H.;
RT Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
RN [6]

SEQUENCE OF 340-564 FROM N.A.
RA Cao J.; Fan W.; Zheng S.;
RT "Genomic analysis of a novel human serine protease SNC19.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
RN [7]

CHARACTERIZATION
RC TISSUE=Milk;
RY MEDLINE=993582; PubMed=10373425;
RA Lin C.Y.; Anders J.; Johnson M.; Dickson R.B.;
RT "Purification and characterization of a complex containing matrix metalloprotease
and a Kunitz-type serine protease inhibitor from human milk.";
RL J. Biol. Chem. 274:18237-18242(1999);
CC -I- FUNCTION: DEGRADES EXTRACELLULAR MATRIX. PROPOSED TO PLAY A ROLE
IN BREAST CANCER INVASION AND METASTASIS. EXHIBITS TRYPSIN-LIKE
ACTIVITY AS DEFINED BY CLEAVAGE OF SYNTHETIC SUBSTRATES WITH ARG
OR LYS AS THE P1 SITE.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
CC -I- SIMILARITY: CONTAINS 4 LDL-RECEPTOR CLASS A DOMAINS.
CC -I- SIMILARITY: CONTAINS 2 CUB DOMAINS.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
use by non-profit institutions. There are no restrictions on its
modified and this statement is as long as its content is in no way
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC

DR EMBL; AF118224; AAD42765.;
DR EMBL; AF133086; AAF00109.;
DR EMBL; AF057145; AAG15395.;
DR EMBL; BC005826; AA05826.;
DR HSSP; P00763; IDPO.
DR GenBank; HGNC; HGNC11344; ST14.
DR MIM; 606797.;
DR MEROPS; S01; 302.;
DR InterPro; IPR000859; CUB-domain.
DR InterPro; IPR001314; Chymotrypsin.

DR InterPro; IPR002172; LDL_recept_A.
DR InterPro; IPR001254; Ser_protease_Try.
DR Pfam; PF00057; ldl_recept_A.;
DR Pfam; PF00431; CUB.;
DR PRINTS; PRO0722; CHYMOTRYPSIN.
DR PRINTS; PRO0261; LDLRECEPTOR.
DR SMART; SM00042; CUB.;
DR SMART; SM00192; LDLa.;
DR SMART; SM00020; TRYPSIN_SPC.;
DR PROSITE; PS01180; CUB.;
DR PROSITE; PS01209; LDLRA.;
DR PROSITE; PS50068; LDLRA.;
DR PROSITE; PS50240; TRYPSIN_DOM.;
DR PROSITE; PS00134; TRYPSIN_HIS.;
DR PROSITE; PS00155; TRYPSIN_SER.;
KW Signal-anchor; Glycoprotein; Hydrolase; Serine protease;
KW Transmembrane; Repeat.

FT DOMAIN 1 55
FT DOMAIN 56 76

FT DOMAIN 77 855
FT DOMAIN 234 334
FT DOMAIN 340 447
FT DOMAIN 452 487
FT DOMAIN 487 524
FT DOMAIN 524 560
FT DOMAIN 566 603
FT DOMAIN 615 654
FT ACT SITE 656 656
FT ACT SITE 711 711
FT ACT SITE 805 805
FT CARBOHYD 109 109
FT CARBOHYD 302 302
FT CARBOHYD 485 485
FT CARBOHYD 772 772
FT CONFLICT 327 329
FT CONFLICT 381 381
FT CONFLICT 674 674
SQ SEQUENCE 855 AA; 94769 MW; 26143132C01F99C9 CRC64;

Query Match 62:2%; Score 51; DB 1; Length 855;
Best Local Similarity 62.5%; Pred. No. 0.33;
Matches 10; Conservative 3; Mi-smatches 3; Indels 0; Gaps 0;

QY 1 VVGGTRAAQEEFEPMV 16
Db 615 VVGGTDADEGEWPQV 630

RESULT 12
TRY_PENNO STANDARD; PRT;
ID TRY_PENNO STANDARD; PRT;
AC P35550;
DT 01-0-FB-1994 (Rel. 28, Created)
DT 01-0-FB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DT Trypsin (EC 3.4.21.4) (Fragment).
OS Penaeus monodon (Penaeid shrimp).
OS Bukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Malacostraca; Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata;
OC Penaeoidea; Penaeidae; Penaeus.
NCBI_TaxID=668;
RN [1]

SEQUENCE;
RC TISSUE=Midgut; PubMed=1963309;
RX MEDLINE=9115186; PubMed=1963309;
RA Lu P.J.; Liu H.C.; Tsai I.H.;
RT "The midgut trypsin of shrimp (Penaeus monodon). High efficiency
toward native protein substrates including collagens.";
RL Biol. Chem. Hoppe-Seyler 371:851-859 (1990).
CC -I- FUNCTION: HIGH EFFICIENCY TOWARD NATIVE PROTEIN SUBSTRATES
INCLUDING COLLAGENS

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.
 CC -!- SUBCELLULAR LOCATION: Extracellular.
 CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
 DR PIR; S11537; S11537.

DR MPROPS; S01_111; -.

DR InterPro; IPR00154; Ser_protease_Try.

DR PIR; PF00089; trypsin; 1.

DR PROSITE; PS001240; TRYPSIN_DOM; PARTIAL.

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.

DR Hydrolase; Serine protease.

FT KW Non_TER; 32; 32 AA; 3330 MW; 96A3A5DF4B19182 CRC64;

SEQUENCE 32 AA; Score 61.0%; DB 1; Length 32;

est Local Similarity 50.0%; Pred. No. 0.016; Indels 0; Gaps 0;

Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy 1 IVGGTRAAQGEPPFMV 16

:||||| :|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

DR PIR; S01_111; -.

DR MPROPS; S01_111; -.

DR InterPro; IPR00154; Ser_protease_Try.

DR PIR; PF00089; trypsin; 1.

DR PROSITE; PS001240; TRYPSIN_DOM; PARTIAL.

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.

DR Hydrolase; Serine protease.

FT KW Non_TER; 32; 32 AA; 3330 MW; 96A3A5DF4B19182 CRC64;

SEQUENCE 32 AA; Score 61.0%; DB 1; Length 32;

est Local Similarity 50.0%; Pred. No. 0.012; Indels 3; Mismatches 5; Indels 0; Gaps 0;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGTAAQGEPPFMV 16

:||||| :|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

DR PIR; S01_111; -.

DR MPROPS; S01_111; -.

DR InterPro; IPR00154; Ser_protease_Try.

DR PIR; PF00089; trypsin; 1.

DR PROSITE; PS001240; TRYPSIN_DOM; PARTIAL.

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.

DR Hydrolase; Serine protease.

FT KW Non_TER; 32; 32 AA; 3330 MW; 96A3A5DF4B19182 CRC64;

SEQUENCE 32 AA; Score 61.0%; DB 1; Length 32;

est Local Similarity 50.0%; Pred. No. 0.012; Indels 3; Mismatches 5; Indels 0; Gaps 0;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGTAAQGEPPFMV 16

:||||| :|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

CC -!- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Xaa, Lys-|-Xaa.

CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.

DR PIR; S01_111; -.

DR MPROPS; S01_111; -.

DR InterPro; IPR00154; Ser_protease_Try.

DR PIR; PF00089; trypsin; 1.

DR PROSITE; PS001240; TRYPSIN_DOM; PARTIAL.

DR PROSITE; PS00134; TRYPSIN_HIS; PARTIAL.

DR PROSITE; PS00135; TRYPSIN_SER; PARTIAL.

DR Hydrolase; Serine protease.

FT KW Non_TER; 32; 32 AA; 3330 MW; 96A3A5DF4B19182 CRC64;

SEQUENCE 32 AA; Score 61.0%; DB 1; Length 32;

est Local Similarity 50.0%; Pred. No. 0.012; Indels 3; Mismatches 5; Indels 0; Gaps 0;

Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 IVGGTAAQGEPPFMV 16

:||||| :|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

Query Match 1 IVGGTAAQGEPPFMV 16

:|:|:|:|:|:|

1 IVGGTAAVTPGEFPYQL 16

:|:|:|:|:|:|

OX	RN	[1] NCBI_TaxID=9986;
	RP	SEQUENCE FROM N.A.
	RC	STRAT=New Zealand white; TISSUE=Testis;
	RX	MEDLINE=94368865; PubMed=808668;
	RA	Richardson R.T., O'Rand M.G.
	RA	"Cloning and sequencing of cDNAs for rabbit preproacrosin and a novel preproacrosin-related cDNA.",
	RT	Biochim. Biophys. Acta 1219:215-218 (1994).
	RT	IT IS THE MAJOR PROTEASE OF MAMMALIAN SPERMATOZOA.
	RL	IT IS A SERINE PROTEASE OF TRYPSIN-LIKE CLEAVAGE SPECIFICITY, IT IS SYNTHESIZED IN A ZYMOGEN FORM, PROACROSIN AND STORED IN THE ACROSMOE.
	CC	-1- CATALYTIC ACTIVITY: HYDROLYSIS OF ARG- AND LYS-BONDS; PREFERENTIAL CLEAVAGE ARG-XAA > LYS-LYS > LYS-XAA.
	CC	-!- SUBUNIT: HEAVY CHAIN (CATALYTIC) AND A LIGHT CHAIN LINKED BY TWO DISULFIDE BONDS.
	CC	-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1.
	CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation of the European Bioinformatics Institute. There are no restrictions on use by non-profit institutions as long as its content is in no modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/annou or send an email to license@isb-sib.ch).
	DR	EMBL; U05204; AAA61630.1; -.
	DR	HSSP; P00763; IDPO.
	DR	MEROPS; S01.223; -.
	DR	InterPro; IPR001314; Chymotrypsin.
	DR	InterPro; IPR001254; Ser protease_Try.
	DR	Pfam; PF00089; trypsin; 1.
	DR	PRINTS; PRO0722; CHMOTRYP SIN.
	DR	SMART; SMD0020; TRYPCOPE.
	DR	PROSITE; PS50240; TRYPSIN_DOM; 1.
	DR	PROSITE; PS00134; TRYPSIN_HIS; 1.
	DR	PROSITE; PS00135; TRYPSIN_SER; 1.
	KW	Hydrolase; Serine protease; Glycoprotein; Zymogen; Sperm; Signal.
	PT	SIGNAL
	FT	16
	FT	CHAIN
	FT	17
	FT	CHAIN
	FT	17
	FT	ACROSIN LIGHT CHAIN (BY SIMILARITY).
	FT	CHAIN
	FT	40
	FT	?
	FT	431
	FT	PROPEP
	FT	?
	FT	400
	FT	288
	FT	DOMAIN
	FT	40
	FT	22
	FT	152
	FT	DISULFID
	FT	22
	FT	160
	FT	DISULFID
	FT	26
	FT	160
	FT	DISULFID
	FT	71
	FT	87
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	175
	FT	244
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	207
	FT	223
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	234
	FT	264
	FT	INTERCHAIN (BY SIMILARITY).
	FT	DISULFID
	FT	71
	FT	87
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	175
	FT	244
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	207
	FT	223
	FT	BY SIMILARITY.
	FT	DISULFID
	FT	234
	FT	264
	FT	INTERCHAIN (BY SIMILARITY).
	FT	DISULFID
	FT	208
	FT	208
	FT	N-LINKED (GLCNAC. . .) (POTENTIAL).
	FT	CARBONYD
	FT	19
	FT	19
	FT	BY SIMILARITY.
	FT	CARBONYD
	FT	208
	FT	208
	FT	N-LINKED (GLCNAC. . .) (POTENTIAL).
	ACT_SITE	86
	ACT_SITE	86
	ACT_SITE	140
	ACT_SITE	140
	ACT_SITE	238
	ACT_SITE	238
	SEQUENCE	431 AA;
	SEQUENCE	464422 MW;
	SEQUENCE	1C015A4EB0C668 CRC64;
	Query Match	59.8%;
	Best Local Similarity	Score 49; DB 1; Length 431;
	Matches 10;	Fred. No. 0.37; Mismatches 2; Indels 4; Gaps 0;
Qy	1	VGGGTTAAQGGFPFMV 16
Dy	40	WVGGGTTAAQGGFPFMV 55

Search completed: April 11, 2003, 17:57:30
Job time : 12 secs

Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 17:56:10 ; Search time 16 Seconds

Title: US-10-006-223-1
Perfect score: 82
Sequence: 1 VVGTRAAQGEFPFMV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched:

1 number of hits satisfying chosen parameters:
283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

POST-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_73;*
1: piri:/*
2: piri2:/*
3: piri3:/*
4: piri:/*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	82	100.0	259	1	TRSMG	trypsin (EC 3.4.21)
2	60.5	73.8	18	2	A61577	24k serine protein trypsin-like protease
3	60	73.2	254	2	S49329	trypsin homolog - alkaline trypsin-like membrane-bound arg membrane type-seri
4	55	67.1	281	2	T13596	trypsin (EC 3.4.21)
5	54	65.9	24	2	PN0655	chymotrypsin (EC 3.4.21)
6	53	64.6	855	2	JC7731	trypsin homolog - alkaline trypsin-like membrane-bound arg membrane type-seri
7	53	64.6	855	2	JC7735	trypsin (EC 3.4.21)
8	52	63.4	20	2	A61327	chymotrypsin (EC 3.4.21)
9	52	63.4	216	1	KYH2O	trypsin (EC 3.4.21)
10	52	63.4	237	1	TRCY1	chymotrypsin-like protease
11	52	63.4	254	2	S35585	trypsin homolog - alkaline trypsin-like membrane-bound arg membrane type-seri
12	52	63.4	274	2	I47078	trypsin (EC 3.4.21)
13	52	63.4	548	2	D82275	trypsin homolog - alkaline trypsin-like membrane-bound arg membrane type-seri
14	51	62.2	638	2	KQMSPL	trypsin (EC 3.4.21)
15	50	61.0	32	2	S11537	chymotrypsin (EC 3.4.21)
16	49	59.8	219	1	TRPGAZ	trypsin (EC 3.4.21)
17	49	59.8	218	1	KYH2C	trypsin (EC 3.4.21)
18	49	59.8	219	1	S47538	trypsin (EC 3.4.21)
19	49	59.8	431	2	S15468	trypsin (EC 3.4.21)
20	49	59.8	613	2	KFHU1	trypsin (EC 3.4.21)
21	49	59.8	625	1	D34817	trypsin (EC 3.4.21)
22	48	58.5	20	2	KQHUP	trypsin (EC 3.4.21)
23	48	58.5	638	1	KQRTPPL	trypsin (EC 3.4.21)
24	48	58.5	638	2	S02176	trypsin (EC 3.4.21)
25	47	57.3	37	2	S52146	trypsin (EC 3.4.21)
26	47	57.3	266	1	KFBO	trypsin (EC 3.4.21)
27	47	57.3	416	1	B87580	trypsin (EC 3.4.21)
28	47	57.3	819	2	I46580	trypsin (EC 3.4.21)
29	46	56.1	275	1	C1HURB	trypsin (EC 3.4.21)

factor IX - rabbit complement subcomponent	3.0	46	56.1	2	146712
collagenolytic protein	3.1	46	56.1	2	B34817
collagenolytic protein	3.2	45	54.9	20	C34817
trypsin (EC 3.4.21)	3.3	45	54.9	42	B34817
signal sequence peptidase (E	3.4	45	54.9	189	F63473
leukocyte elastase	3.5	45	54.9	267	ELHDLL
coagulation factor	3.6	45	54.9	285	I48144
coagulation factor	3.7	45	54.9	2	A30351
coagulation factor	3.8	45	54.9	452	1
coagulation factor	3.9	45	54.9	461	1
complement factor	4.0	44	53.7	583	KFHU
complement factor	4.1	44	53.7	1019	A29154
enteropeptidase (E	4.2	43	52.4	250	S55493
serine proteinase	4.3	43	52.4	251	TRHD2
azurocidin	4.4	43	52.4	415	A34170
acrosin (EC 3.4.21)	4.5	43	52.4	421	S11674
acrosin (EC 3.4.21)					

ALIGNMENTS

RESULT 1	TRSMG	trypsin (EC 3.4.21.4) precursor - Streptomyces griseus
C;Species:	Streptomyces griseus	
C;Accession:	ATCC10137	
C;Date:	24-Apr-1984	#sequence_change 07-May-1999
R;Kim, J.C.; Cha, S.H.; Jeong, S.T.; Oh, S.K.; Byun, S.M.		
R;Olaofon, R.W.; Jurasek, L.; Carpenter, M.R.; Smillie, L.B.		
Biochemistry 14, 1168-1171, 1975		
A;Title: Molecular cloning and nucleotide sequence of Streptomyces griseus trypsin gene		
A;Reference number: JQ1302; MUID:92095977; PMID:1755852		
A;Molecule type: DNA		
A;Accession: A00962		
A;Cross-references: GB: M64471		
A;Experimental source: strain ATCC10137		
R;Olaofon, R.W.; Jurasek, L.; Carpenter, M.R.; Smillie, L.B.		
Biochemistry 14, 1168-1171, 1975		
A;Title: Amino acid sequence of Streptomyces griseus trypsin. Cyanogen bromide fragment:		
A;Residues: 1-259		
C;Molecule type: protein		
A;Accession: A00962		
A;Residues: 37-95, 98-259 <OLA>		
R;Read, R.J.; James, M.N.G.		
J. Mol. Biol. 200, 523, 1988		
A;Title: Refined crystal structure of Streptomyces griseus trypsin at 1.7 angstroms resolution		
A;Reference number: A44574; MUID:3135412		
A;Contents: annotation; x-ray crystallography; status experimental		
A;Note: residues 96-97 modelled as Gly-Ala		
C;Genetics:		
A;Gene: sprt		
C;Superfamily: trypsin; trypsin homology		
C;Keywords: hydrolase; serine proteinase		
F;1-32/Domain: signal sequence #status predicted <SIG>		
F;33-36/Domain: propeptide #status predicted <PRO>		
F;37-258/Domain: trypsin #status experimental		
F;37-252/Domain: trypsin homology <TRY>		
F;73-118/208/Active site: His, Asp, Ser #status experimental		
Query Match Best Local Similarity Score 82; DB 1; Length 259; Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY 1 VGGTRAAQEEFPFMV 16		
Db 37 VGGTRAAQEEFPFMV 52		
RESULT 2		
A61577 24k serine proteinase (EC 3.4.21.-) - Streptomyces fradiae (fragment)		
C;Species: Streptomyces fradiae		

C;Date: 28-Oct-1994 #sequence_revision 28-Oct-1994 #text_change 06-Jan-1995
 C;Accession: A61577
 R;Sinha, U.; Wolz, S.A.; Lad, P.J.
 Int. J. Biochem. 23, 979-984, 1991

A;Title: Two new extracellular serine proteases from *Streptomyces fradiae*.

A;Reference number: A61577; MUID:92155439; PMID:1788859

A;Accession: A61577
 A;Molecule type: protein

A;Residues: 1-18 <SIN>

C;Keywords: extracellular protein; hydrolase; serine proteinase

Query Match 1 VVGCTRAAQGEFPFMV 16
 Best Local Similarity 73.8%; Score 60.5%; DB 2; Length 18;

Qy 1 VVGCTRAAQGEFPFMV 16
 Db 1 VVGCTRAAQ-EFPFMV 15

RESULT 3

54339 trypsin-like proteinase 1 (EC 3.4.21.-) precursor - imperfect fungus (*Metarhizium anisopliae*

C;Species: *Metarhizium anisopliae*
 C;Date: 16-Feb-1995 #sequence_revision 12-May-1995 #text_change 22-Jun-1999

C;Accession: JC4517; S9329

R;Smithson, S.L.; Paterton, I.C.; Bailey, A.M.; Screen, S.E.; Hunt, B.A.; Cobb, B.D.; Co

Gene 166, 161-165, 1995

A;Title: Cloning and characterisation of a gene encoding a cuticle-degrading protease fr

A;Reference number: JC4517; MUID:96105219; PMID:8529882

A;Accession: JC4517
 A;Molecule type: DNA

A;Cross references: EMBL:X78875; NID:9556656; PIDN:CAA55477.1; PID:9556657

A;Experimental source: MB1

A;Note: submitted to the EMBL Data Library, April 1994

C;Genetics:

A;Gene: try1

A;Introns: 86/3; 149/2

C;Superfamily: trypsin; trypsin homology

C;Keywords: endoplasmic reticulum; hydrolase; serine proteinase

C;21/Domain: signal sequence #status predicted <SIG>

C;29/Domain: propeptide #status predicted <PRP>

C;30-254/Domain: trypsin-like proteinase #status predicted <MAT>

C;70,112,209/Active Site: His, Asp, Ser #status predicted

C;70,112,209/Domain: trypsin homology <TRY>

Query Match 1 VVGCTRAAQGEFPFMV 16
 Best Local Similarity 73.2%; Score 60; DB 2; Length 254;

Qy 1 VVGCTRAAQGEFPFMV 16
 Db 30 IVGGSPAAGAEFPFIV 45

RESULT 4

T13596 trypsin homolog - fruit fly (*Drosophila melanogaster*)

N;Alternate names: protein EG80H7.1

C;Species: *Drosophila melanogaster*

C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C;Accession: T13596

R;Benois, P.

A;Description: Sequencing the distal X chromosome of *Drosophila melanogaster*.

A;Reference number: Z176667

A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA

A;Residues: 1-281 <BEN>

A;Cross-references: EMBL:AL031027; PIDN:CAA19843.2

C;Genetics:

A;Cross-references: FlyBase:FBgn0000481

A;Introns: 159/3

A;Note: EG80H7.1

C;Superfamily: trypsin; trypsin homology

Query Match 1 VVGCTRAAQGEFPFMV 16
 Best Local Similarity 67.1%; Score 55; DB 2; Length 281;

Qy 1 VVGCTRAAQGEFPFMV 16
 Db 27 IVNGTTAGPGSEFPFV 42

RESULT 5

PN0655 alkaline trypsin-like serine proteinase (EC 3.4.21.-) F-II - earthworm (*Lumbricus rubellus*)

N;Alternate names: fibrinolytic enzyme

C;Species: *Lumbricus rubellus* (humus earthworm)

C;Accession: 14-Jul-1994 #sequence_revision 14-Jul-1994 #text_change 07-May-1999

C;Accession: PN0655

R;Nakajima, N.; Miura, H.; Sumi, H.

Biosci. Biotechnol. Biochem. 57, 1726-1730, 1993

A;Title: Characterization of potent fibrinolytic enzymes in earthworm, *Lumbricus rubellus*

A;Reference number: PN0653; MUID:9408032; PMID:7764268

A;Molecule type: Protein

A;Residues: 1-24 <NAK>

C;Comment: This enzyme is a stable and potent fibrinolytic enzyme and is absorbed from

C;Keywords: hydrolase; serine proteinase

Query Match 1 VVGCTRAAQGEFPFMV 16
 Best Local Similarity 56.2%; Score 54; DB 2; Length 24;

Qy 1 VVGCTRAAQGEFPFMV 16
 Db 1 VIGGTNNSPGEPFWQL 16

RESULT 6

JC7731 membrane-bound arginine-specific serine proteinase precursor - rat

C;Species: *Rattus norvegicus* (Norway rat)

C;Accession: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 14-Dec-2001

R;Kishi, K.; Yamazaki, K.; Yasuda, I.; Yahagi, N.; Ichinose, M.; Tsuchiya, Y.; Athauda, J. Biochem. 130, 425-430, 2001

A;Title: Characterization of a membrane-bound arginine-specific serine protease from rat

A;Reference number: JC7731; MUID:21421307; PMID:11530019

A;Accession: JC7731

A;Molecule type: mRNA

A;Residues: 1-855 <KIS>

A;Cross-references: DDBJ:AB049169

A;Experimental source: strain Male, 7-week-old

C;Comment: This enzyme, localized mainly on brushborder membranes of the intestine, parti

C;Keywords: protein digestion

Query Match 1 VVGCTRAAQGEFPFMV 16
 Best Local Similarity 62.5%; Score 53; DB 2; Length 855;

Qy 1 VVGCTRAAQGEFPFMV 16
 Db 615 VVGGTNADEGENFWQV 630

RESULT 7

JC7775 membrane type-serine protease 1 - rat

C;Species: *Rattus norvegicus* (Norway rat)

C;Accession: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002

C;Accession: JC7775


```

Query Match      62.2%; Score 51; DB 1; Length 638;
Best Local Similarity 56.2%; Pred. No. 0.91;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy   1 VVGGTRAAQCEFPFPMV 16
      :|||:|||:|||:|||:
Db   391 IVGGTNASLGEWPWQV 406

RESULT 15
S11537
trypsin (BC 3.4.21.4) 1 - penaeid shrimp (Penaeus monodon) (fragment)
C;Species: Penaeus monodon
C;Date: 21-Nov-1993 #sequence_revision 19-Jul-1996 #text_change 16-Feb-1997
C;Accession: S11537
R;Lu, P.J.; Liu, H.C.; Tsai, I.H.
Biol. Chem. Hoppe-Seyler 371, 851-859, 1990
A;Title: The midgut trypsin of shrimp (Penaeus monodon). High efficiency toward native
A;Reference number: S11537; MUID:9115186; PMID:1963309
A;Accession: S11537
A;Molecule type: protein
A;Residues: 1-32 < LUA>
C;Superfamily: trypsin; trypsin homology
C;Keywords: hydrolase; protein digestion; serine proteinase

Query Match      61.0%; Score 50; DB 2; Length 32;
Best Local Similarity 50.0%; Pred. No. 0.068;
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Qy   1 VVGGTRAAQCEFPFPMV 16
      :|||:|||:|||:|||:
Db   1 IVGGTAVTPGEFPYQL 16

Search completed: April 11, 2003, 17:58:28
Job time : 17 secs

```

RESULT 14
KOMSPL
 plasma kallikrein (EC 3.4.21.34) precursor - mouse
 C;species: *Mus musculus* (house mouse)
 C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 18-Jun-1999
 C;Accession: A36557
 R;Seidell, N.G.; Sawyer, N.; Hamelin, J.; Mion, P.; Beaubien, G.; Brachypa, L.; Rochemon
 DNA Cell Biol. 9:737-748, 1990
 A;Title: Mouse plasma kallikrein: cDNA structure, enzyme characterization, and comparison
 A;Reference number: A36557; MUID:91090844; PMID:2264928
 A;Accession: A36557
 A;Molecule type: mRNA
 A;Residues: 1-638 <SEI>
 A;Cross-references: GB:MS58588; PID:AAA63393.1; PID:9200359
 A;Note: part of this sequence, including the amino ends of both the heavy and light chain
 C;Comment: This protein, synthesized in the liver, circulates as a noncovalent complex w
 C;Comment: The zymogen is activated by factor XIIa, which cleaves the molecule into a 11

Copyright GenCore version 5.1.4_p5_4578
(c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw mode1

Run on: April 11, 2003, 17:57:21 ; Search time 14 Seconds
(without alignments)
69.870 Million cell updates/sec

Title: US-10-006-223-1
Perfect score: 82
Sequence: 1 VVGITRAAQGEEPFMV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 248812 seqs, 61136040 residues

11 number of hits satisfying chosen parameters: 248812

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications AA.*

```
1: /cgn2_6/ptodata/2/pubbaa/US08_NEW_PUB.PEP/*
2: /cgn2_6/ptodata/2/pubbaa/PCT_NEW_PUB.PEP/*
3: /cgn2_6/ptodata/2/pubbaa/US06_NEW_PUB.PEP/*
4: /cgn2_6/ptodata/2/pubbaa/US06_PUBCOMB.PEP/*
5: /cgn2_6/ptodata/2/pubbaa/US07_NEW_PUB.PEP/*
6: /cgn2_6/ptodata/2/pubbaa/US07_PUBCOMB.PEP/*
7: /cgn2_6/ptodata/2/pubbaa/US08_PUBCOMB.PEP/*
8: /cgn2_6/ptodata/2/pubbaa/US08_PUBCOMB.PEP/*
9: /cgn2_6/ptodata/2/pubbaa/US09_NEW_PUB.PEP/*
10: /cgn2_6/ptodata/2/pubbaa/US09_PUBCOMB.PEP/*
11: /cgn2_6/ptodata/2/pubbaa/US10_NEW_PUB.PEP/*
12: /cgn2_6/ptodata/2/pubbaa/US10_PUBCOMB.PEP/*
13: /cgn2_6/ptodata/2/pubbaa/US16_NEW_PUB.PEP/*
14: /cgn2_6/ptodata/2/pubbaa/US60_PUBCOMB.PEP/*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	53	64.6	855	10	US-09-900-751-2		Sequence 2, Appli
2	52	63.4	20	9	US-09-938-269-13		Sequence 13, Appli
3	51	62.2	241	9	US-10-095-700A-4		Sequence 4, Appli
4	51	62.2	241	9	US-10-092-000A-2		Sequence 2, Appli
5	51	62.2	620	10	US-09-925-301-1193		Sequence 1193, Appli
6	51	62.0	855	25	US-09-938-269-4		Sequence 2, Appli
7	50	61.0	25	9	US-09-808-602-104		Sequence 4, Appli
8	49	59.8	625	9	US-09-938-269-12		Sequence 104, Appli
9	48	58.5	20	9	US-09-808-602-102		Sequence 12, Appli
10	48	58.5	638	9	US-09-808-602-102		Sequence 102, Appli
11	48	58.5	1128	10	US-09-988-615-97		Sequence 97, Appli
12	46	56.1	449	10	US-09-935-301-612		Sequence 612, Appli
13	46	56.1	688	10	US-09-974-198-7		Sequence 7, Appli
14	46	56.1	688	10	US-09-874-238-7		Sequence 94, Appli
15	46	56.1	705	9	US-09-808-602-94		Sequence 2, Appli
16	46	54.9	18	10	US-09-861-708-2		Sequence 10, Appli
17	45	54.9	20	9	US-09-938-269-10		Sequence 11, Appli
18	45	54.9	20	9	US-09-938-269-11		Sequence 3, Appli
19	45	54.9	218	10	US-09-861-708-3		

ALIGNMENTS

```
RESULT 1
US-09-900-751-2
; Sequence 2, Application US/09900751
; Patent No. US2002006653A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Keith D.
; ATTORNEY: Levitin, Michael W.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING SERINE
; TITLE OF INVENTION: PROTEASE GENE DISRUPTIONS
; FILE REFERENCE: R-386
; CURRENT APPLICATION NUMBER: US/09/900,751
; CURRENT FILING DATE: 2001-07-06
; PRIOR APPLICATION NUMBER: US 60/217,449
; PRIOR FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: US 60/223,170
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 60/223,460
; PRIOR FILING DATE: 2000-08-07
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-900-751-2

Query Match 64.6%; Score 53; DB 10; Length 855;
Best Local Similarity 62.5%; Pred. No. 0.67;
Matches 10; Conservative 3; Missmatches 3; Indels 0; Gaps 0;

Qy 1 VVGITRAAQGEEPFMV 16
Db 615 VVGITRAAQGEEPFMV 630

RESULT 2
US-09-938-269-13
; Sequence 13, Application US/09938269
; Publication No. US2003000751A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; ATTORNEY: Cowling, Didier S.P.
```

APPLICANT: Hubbel, Jeffrey A.
 APPLICANT: van de Wetering, Petra
 TITLE OF INVENTION: Treatment of Trauma
 FILE REFERENCE: 314512-105
 CURRENT APPLICATION NUMBER: US/09/938,269
 CURRENT FILING DATE: 2001-08-23
 NUMBER OF SEQ ID NOS: 17
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 13
 LENGTH: 20
 TYPE: PRT
 ORGANISM: Crayfish protease
 US-09-938-269-13

Query Match 63.4%; Score 52; DB 9; Length 20;
 Best Local Similarity 56.2%; Pred. No. 0.019; 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEPPMV 16
 :|||||:|||||:
 1 VVGTTDADGEWPWQV 16

RESULT 3
 US-10-099-700A-4
 Sequence 4, Application US/10099700A
 Publication No. US200300050372A1
 GENERAL INFORMATION:
 APPLICANT: Edwin L. Madison
 APPLICANT: Edgar O. Ong
 TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING TRANSMEMBRANE SERINE PROTEASE 7,
 TITLE OF INVENTION: ENCODED POLYPEPTIDES AND METHODS BASED THEREON
 FILE REFERENCE: 2A74-1613
 CURRENT APPLICATION NUMBER: US/10/099,700A
 CURRENT FILING DATE: 2002-05-24
 PRIOR APPLICATION NUMBER: 60/275,592
 PRIOR FILING DATE: 2001-03-13
 NUMBER OF SEQ ID NOS: 22
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 4
 LENGTH: 241
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-099-700A-4

Query Match 62.2%; Score 51; DB 9; Length 241;
 Best Local Similarity 62.5%; Pred. No. 0.38; 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEPPMV 16
 :|||||:|||||:
 1 VVGTTDADGEWPWQV 16

RESULT 4
 US-10-092-004A-2
 Sequence 2, Application US/10092004A
 Publication No. US20030050251A1
 GENERAL INFORMATION:
 APPLICANT: Corvas International, Inc.
 APPLICANT: Madison, Edwin L.
 APPLICANT: Semple, Joseph Edward
 APPLICANT: Coombs, Gary Samuel
 APPLICANT: Reiner, John Eugene
 APPLICANT: Ong, Edgar O.
 APPLICANT: Araldi, Gian Luca
 TITLE OF INVENTION: Inhibitors of Serine Protease Activity of Matrixinase or MTSP1
 FILE REFERENCE: 018813/0282105
 CURRENT FILING DATE: 2002-09-12
 PRIOR APPLICATION NUMBER: PCT/US01/28137
 PRIOR FILING DATE: 2001-09-07
 PRIOR APPLICATION NUMBER: 09/657,986

Query Match 62.2%; Score 51; DB 9; Length 241;
 Best Local Similarity 62.5%; Pred. No. 0.38; 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEPPMV 16
 :|||||:|||||:
 1 VVGTTDADGEWPWQV 490

RESULT 5
 US-09-325-301-1193
 Sequence 1193, Application US/09925301
 Patent No. US20020052308A1
 GENERAL INFORMATION:
 APPLICANT: Rosen, et al.
 TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 FILE REFERENCE: PA106
 CURRENT APPLICATION NUMBER: US/09/925,301
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: PCT/US00/05882
 PRIOR FILING DATE: 2000-03-08
 PRIOR APPLICATION NUMBER: 60/124,270
 PRIOR FILING DATE: 1999-03-12
 NUMBER OF SEQ ID NOS: 1694
 SEQ ID NO: 1193
 LENGTH: 620
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: SITE
 LOCATION: (375)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (501)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (532)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 NAME/KEY: SITE
 LOCATION: (546)
 OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
 US-09-925-301-1193

Query Match 62.2%; Score 51; DB 10; Length 620;
 Best Local Similarity 62.5%; Pred. No. 1; 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEPPMV 16
 :|||||:|||||:
 1 VVGTTDADGEWPWQV 490

RESULT 6
 US-10-099-700A-2
 Sequence 2, Application US/10099700A
 Publication No. US2003008372A1
 GENERAL INFORMATION:
 APPLICANT: Edwin L. Madison
 APPLICANT: Edgar O. Ong
 TITLE OF INVENTION: NUCLEAR ACID MOLECULES ENCODING TRANSMEMBRANE SERINE PROTEASE 7,
 FILE REFERENCE: 24745-1613
 CURRENT FILING DATE: 2002-09-12
 PRIOR APPLICATION NUMBER: PCT/US01/28137
 PRIOR FILING DATE: 2001-09-07
 PRIOR APPLICATION NUMBER: US/10/099,700A
 CURRENT APPLICATION NUMBER: US/10/099,700A

```

; CURRENT FILING DATE: 2002-05-24
; PRIOR APPLICATION NUMBER: 60/275,592
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 855
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-099-700A-2

Query Match 62.2%; Score 51; DB 9; Length 855;
Best Local Similarity 62.5%; Pred. No. 1.4;
Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGGTRAAGGEPFPVW 16
Db 615 VVGGTDADEGENFWQV 630

RESULT 7
US-09-938-269-4
; Sequence 4, Application US/09938269
; Publication No. US2003007951A1
; GENERAL INFORMATION:
; APPLICANT: Franklin, Richard L.
; APPLICANT: Cowling, Didier S.P.
; APPLICANT: Hubbel, Jeffrey A.
; APPLICANT: van de Wetering, Petra
; TITLE OF INVENTION: Treatment of Trauma
; FILE REFERENCE: 3:14572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Kamchatka crab
; FEATURE:
; NAME KEY: VARIANT
; LOCATION: (1) .. (20)
; OTHER INFORMATION: xaa = Any Amino Acid
US-09-938-269-12

Query Match 58.5%; Score 48; DB 9; Length 20;
Best Local Similarity 50.0%; Pred. No. 0.087;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGGTRAAGGEPFPV 14
Db 1 IVGGSEATSGQFPY 14

RESULT 8
US-09-808-602-104
; Sequence 104, Application US/09808602
; Patent No. US2002015511A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015511A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15:966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 104
; LENGTH: 625
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-104

Query Match 59.8%; Score 49; DB 9; Length 625;
Best Local Similarity 50.0%; Pred. No. 2.2;
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGGTRAAGGEPFPVW 16
Db 388 IVGGTASVRGEWPHQV 403

RESULT 9
US-09-938-269-12
; Sequence 12, Application US/09938269
; Publication No. US2003007951A1
; GENERAL INFORMATION:
; APPLICANT: Cowling, Didier L.
; APPLICANT: Hubbel, Jeffrey A.
; APPLICANT: van de Wetering, Petra
; TITLE OF INVENTION: Treatment of Trauma
; FILE REFERENCE: 3:14572-105
; CURRENT APPLICATION NUMBER: US/09/938,269
; CURRENT FILING DATE: 2001-08-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 12
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Kamchatka crab
; FEATURE:
; NAME KEY: VARIANT
; LOCATION: (1) .. (20)
; OTHER INFORMATION: xaa = Any Amino Acid
US-09-938-269-12

Query Match 58.5%; Score 48; DB 9; Length 20;
Best Local Similarity 50.0%; Pred. No. 0.087;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGGTRAAGGEPFPV 14
Db 1 IVGGSEATSGQFPY 14

RESULT 10
US-09-808-602-102
; Sequence 102, Application US/09808602
; Patent No. US2002015511A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US2002015511A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15:966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102

```

; LENGTH: 638
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-888-615-97
 ; Sequence 97, Application US/09888615
 ; Patent No. US20020061856A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PLOOMAN, GREGORY
 ; APPLICANT: WHYTE, DAVID
 ; APPLICANT: CAENBEEPEI, SEAN
 ; APPLICANT: CHARYDCZAK, GLEN
 ; APPLICANT: MANNING, GERRARD
 ; APPLICANT: SUDARSANAM, SUCHA
 ; TITLE OF INVENTION: NOVEL PROTEASES
 ; FILE REFERENCE: 038602/1214
 ; CURRENT APPLICATION NUMBER: US/09/888, 615
 ; CURRENT FILING DATE: 2001-06-26
 ; PRIOR APPLICATION NUMBER: 60/214, 047
 ; PRIOR FILING DATE: 2000-06-26
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 97
 ; LENGTH: 1128
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 *US-09-888-615-97

; LENGTH: 638
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-808-602-102
 ; Sequence 98, Application US/09925302
 ; Patent No. US2002004494A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
 ; FILE REFERENCE: PA104
 ; CURRENT APPLICATION NUMBER: US/09/925, 302
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: PCT/US00/05918
 ; PRIOR FILING DATE: 2000-03-08
 ; PRIOR APPLICATION NUMBER: 60/124, 270
 ; PRIOR FILING DATE: 1999-03-12
 ; NUMBER OF SEQ ID NOS: 896
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO: 612
 ; LENGTH: 449
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (117)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 *-

; LOCATION: (284)
 ; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
 US-09-925-302-512
 ; Query Match Score 48; DB 9; Length 638;
 Best Local Similarity 50.0%; Pred. No. 3.3%;
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 VVGTRAAQGEFPFMV 16
 Db 391 IVGCTNNSWGEWPWQV 406

; RESULT 13
 US-09-874-198-7
 ; Sequence 7, Application US/09874198
 ; Patent No. US20020082208A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jonsenius, Jens Chr.
 ; APPLICANT: Thiel, Steffen
 ; TITLE OF INVENTION: MASP-2 COMPLEMENT-FIXING ENZYME, AND
 ; TITLE OF INVENTION: USES FOR IT
 ; FILE REFERENCE: 09011-002002
 ; CURRENT APPLICATION NUMBER: US/09/874,198
 ; CURRENT FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: 09/054,218
 ; PRIOR FILING DATE: 1998-04-02
 ; PRIOR APPLICATION NUMBER: 60/042,678
 ; PRIOR FILING DATE: 1997-04-03
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 7
 ; LENGTH: 688
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-874-198-7

; Query Match Score 46; DB 10; Length 688;
 Best Local Similarity 43.8%; Pred. No. 5;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 VVGTRAAQGEFPFMV 16
 Db 447 IIGGQKAKMGNGFPWQV 462

; RESULT 14
 US-09-874-238-7
 ; Sequence 7, Application US/09874238
 ; Patent No. US20020082209A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jonsenius, Jens Chr.
 ; APPLICANT: Thiel, Steffen
 ; TITLE OF INVENTION: MASP-2 COMPLEMENT-FIXING ENZYME, AND
 ; TITLE OF INVENTION: USES FOR IT
 ; FILE REFERENCE: 09011-002003
 ; CURRENT APPLICATION NUMBER: US/09/874,238
 ; CURRENT FILING DATE: 2001-06-04
 ; PRIOR APPLICATION NUMBER: 09/054,218
 ; PRIOR FILING DATE: 1998-04-02
 ; PRIOR APPLICATION NUMBER: 60/042,678
 ; PRIOR FILING DATE: 1997-04-03
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO: 7
 ; LENGTH: 688
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-874-238-7

; Query Match Score 46; DB 10; Length 688;
 Best Local Similarity 43.8%; Pred. No. 7.8;
 Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
 Qy 1 VVGTRAAQGEFPFMV 16
 Db 208 IIGGQKAKMGNGFPWQV 223

Qy 1 VVGTRAAQEFPMV 16
Db 447 IIGGQAKMGNFPWQV 462

RESULT 15

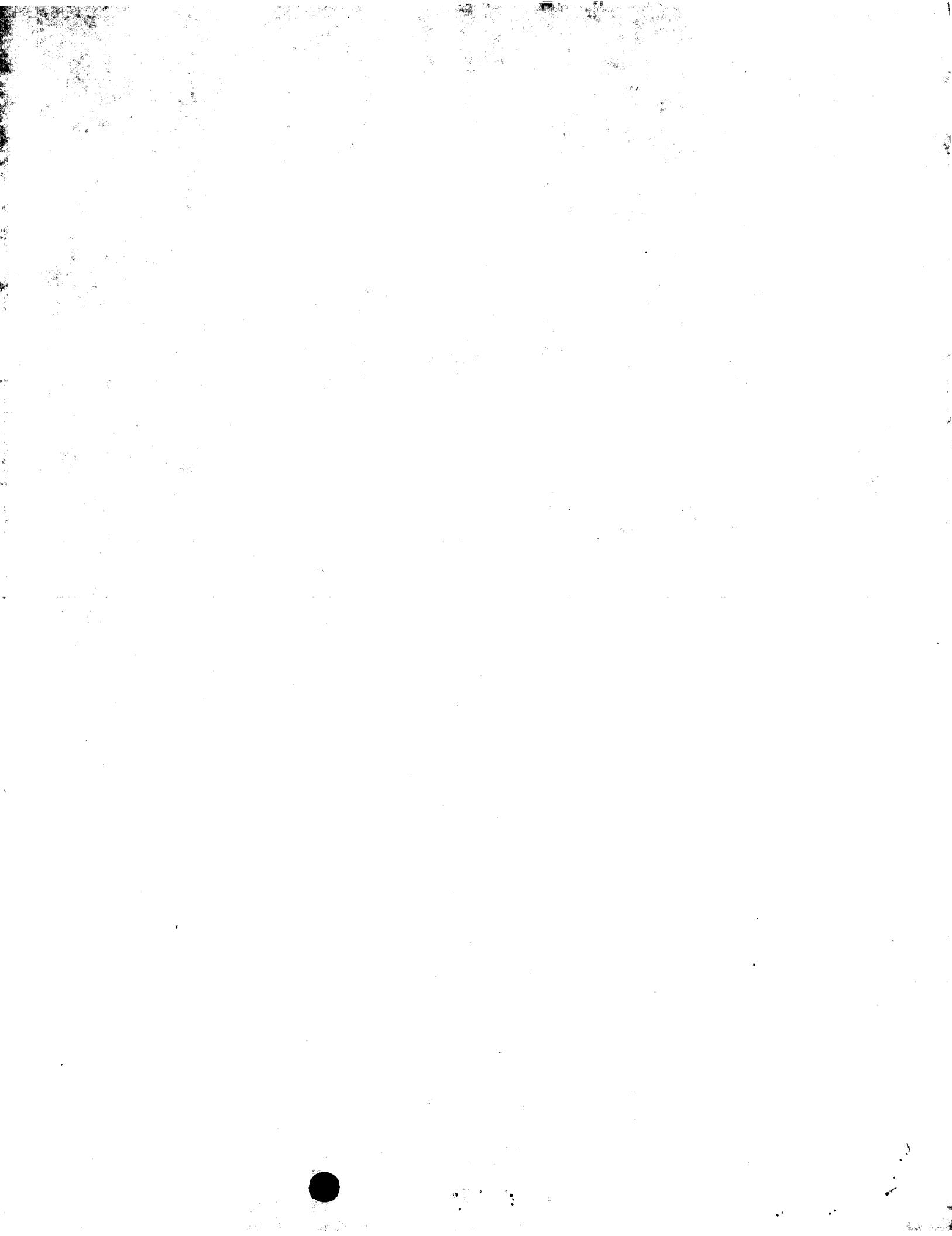
US-09-808-602-94
; Sequence 94, Application US/09808602

; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shimkets, Richard A
; APPLICANT: Herman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: Machougal, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; CURRENT FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 94
; LENGTH: 705
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-808-602-94

Query Match 56.1%; Score 46; DB 9; Length 705;
Best Local Similarity 43.8%; Pred. No. 8;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Qy 1 VVGTRAAQEFPMV 16
Db 464 IIGGQAKMGNFPWQV 479

Search completed: April 11, 2003, 17:59:11
Job time : 15 secs



Copyright GenCore version 5.1.4.D5_4578
(c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 17:56:35 ; Search time 15 Seconds
(without alignments)
31.384 Million cell. updates/sec

Title: US-10-006-223-1
Perfect score: 82
Sequence: 1 VVGCTRRAQGEFPFMV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

1 number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgm2_6_ptodata/1/iaa/5A_COMB.pep:*

2: /cgm2_6_ptodata/1/iaa/5B_COMB.pep:*

3: /cgm2_6_ptodata/1/iaa/6A_COMB.pep:*

4: /cgm2_6_ptodata/1/iaa/6B_COMB.pep:*

5: /cgm2_6_ptodata/1/iaa/PCUTS_COMB.pep:*

6: /cgm2_6_ptodata/1/iaa/backfiles..pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match	Length	DB ID	Description
1	82	100.0	223	1	US-08-278-091-13	Sequence 13, Appli
2	82	100.0	223	1	US-08-483-859-13	Sequence 13, Appli
3	82	100.0	223	2	US-08-472-173-13	Sequence 13, Appli
4	82	100.0	223	2	US-08-487-167-13	Sequence 13, Appli
5	82	100.0	223	2	US-08-482-816-13	Sequence 13, Appli
6	82	100.0	223	2	US-08-295-149-13	Sequence 13, Appli
7	82	100.0	223	2	US-08-801-499-13	Sequence 13, Appli
8	82	100.0	223	2	US-08-615-271-13	Sequence 13, Appli
9	82	100.0	223	3	US-09-074-660-13	Sequence 13, Appli
10	82	100.0	223	3	US-09-074-659-13	Sequence 13, Appli
11	82	100.0	223	3	US-09-104-468-13	Sequence 13, Appli
12	82	100.0	223	4	US-09-104-466-13	Sequence 13, Appli
13	82	100.0	223	4	US-09-106-467-13	Sequence 13, Appli
14	59	72.0	224	1	US-08-553-516-2	Sequence 2, Appli
15	59	72.0	248	1	US-08-923-130-2	Sequence 2, Appli
16	59	72.0	248	2	US-08-923-426-4	Sequence 4, Appli
17	59	72.0	248	3	US-08-814-915-4	Sequence 4, Appli
18	59	72.0	248	5	PCT-US95-07743-4	Sequence 4, Appli
19	53	64.6	902	4	US-09-644-600-10	Sequence 10, Appli
20	52	63.4	20	2	US-08-389-540-15	Sequence 15, Appli
21	52	63.4	20	2	US-08-600-273A-15	Sequence 15, Appli
22	52	63.4	20	3	US-08-486-820-15	Sequence 15, Appli
23	52	63.4	20	4	US-09-220-731-15	Sequence 15, Appli
24	52	63.4	237	4	US-08-164-919-3	Sequence 3, Appli
25	52	63.4	237	5	PCT-US94-14073-3	Sequence 3, Appli
26	51	62.2	855	2	US-09-022-337-2	Sequence 2, Appli
27	51	62.2	855	4	US-09-644-600-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-278-091-13
; Sequence 13, Application US/08278091
; Patent No. 5506139

GENERAL INFORMATION:
; APPLICANT: LOOSMORE, Sheena M
; APPLICANT: YANG, Yan-Ping
; APPLICANT: CHONG, Pele
; APPLICANT: OMEN, Raymond P.
; APPLICANT: KLEIN, Michael H.
TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sim E. McBurney
; STREET: Suite 701, 330 University Avenue
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: MSG 1R7

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRERNT APPLICATION DATA:
; APPLICATION NUMBER: US/08/278,091
; FILING DATE: 21-JUL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Stewart, Michael I
; REGISTRATION NUMBER: 24,973
; REPERENCE/DOCKET NUMBER: 1038-371
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 595-1155
; TELEFAX: (416) 595-1163
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 223 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-278-091-13

Query Match 100.0%; Score 82; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 3.1e-06;
Matches 16; Conservave 0; Mismatches 0; Indels 0;
Gaps 0;
Qy 1 VVGCTRRAQGEFPFMV 16

Db 1 VVGTRAAQGEPPFMV 16

RESULT 2
US-08-483-859-13

; Sequence 13, Application US/08483859

; Patent No. 565636

; GENERAL INFORMATION:

; APPLICANT: LOOSMORE, Sheena M.

; APPLICANT: YANG, Yan-Ping

; APPLICANT: CHONG, Pele

; APPLICANT: Oomen, Raymond P.

; APPLICANT: KLEIN, Michel H.

; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Activity

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: Suite 701, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: MSG 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/296,149

; FILING DATE: 07-JUN-1995

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,859

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 435

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/08/296,149

; FILING DATE: 26-AUG-1994

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US/08/278,091

; FILING DATE: 21-JUL-1994

; ATTORNEY/AGENT INFORMATION:

; NAME: Stewart, Michael I.

; REGISTRATION NUMBER: 24,973

; REFERENCE/DOCKET NUMBER: 1038-493 MIS:vg

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (416) 595-1155

; TELEFAX: (416) 595-1163

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 223 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; US-08-472-173-13

; INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Activity

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: Suite 701, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: MSG 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

RESULT 3
US-08-472-173-13

; Sequence 13, Application US/08472173

; Patent No. 5665353

; GENERAL INFORMATION:

; APPLICANT: LOOSMORE, Sheena M

; APPLICANT: YANG, Yan-Ping

; APPLICANT: CHONG, Pele

; APPLICANT: Oomen, Raymond P.

; APPLICANT: KLEIN, Michel H.

; TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Activity

; NUMBER OF SEQUENCES: 23

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Sim & McBurney

; STREET: Suite 701, 330 University Avenue

; CITY: Toronto

; STATE: Ontario

; COUNTRY: Canada

; ZIP: MSG 1R7

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/487,167
 FILING DATE: 07-JUN-1995
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/296,149
 FILING DATE: 26-AUG-1994
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/278,091
 FILING DATE: 21-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REFERENCE/DOCKET NUMBER: 24, 973
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 FAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-487-167-13

Query Match 100.0%; Score 82; DB 2; Length 223;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEFPFMV 16
 Db 1 VVGTRAAQGEFPFMV 16

RESULT 6
 US-08-296-149-13
 Sequence 13, Application US/08296149
 GENERAL INFORMATION:
 PATENT NO. 593573
 APPLICANT: LOOSMORE, Sheena M
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: OOMEN, Raymond P.
 APPLICANT: KLEIN, Michael H.
 TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Activity
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: Suite 701, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: MSG 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/296,149
 FILING DATE:
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-330
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 FAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-296-149-13

Query Match 100.0%; Score 82; DB 2; Length 223;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 0; Mismatches 0;
 Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEFPFMV 16
 Db 1 VVGTRAAQGEFPFMV 16

RESULT 7

US-08-801-499-13
Sequence 13, Application US/08801499

Patent No. 5962430

GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheena M

APPLICANT: YANG, Yan-Ping

APPLICANT: CHONG, Pele

APPLICANT: OOMEN, Raymond P.

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with Reduced Protease Act

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: MSG 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/615,271

FILING DATE: 20-JUN-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-580

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

8-801-499-13

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/482,816

FILING DATE: 07-JUN-1995

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/296,149

FILING DATE: 26-AUG-1994

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: US 08/278,091

FILING DATE: 21-JUL-1994

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-671 MIS:jb

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

8-801-499-13

Query Match

100.0%

Score 82;

DB 2;

Length 223;

Best Local Similarity 100.0%

Pred. No. 3.1e-06;

Matches 16;

Conservative 0;

MisMatches 0;

Indels 0;

Gaps 0;

Db 1 VYGGTRAAQSEFPFMV 16

Qy 1 VYGGTRAAQSEFPFMV 16

Db 1 VYGGTRAAQSEFPFMV 16

RESULT 8

US-08-615-271-13

Sequence 13, Application US/08615271

Patent No. 5981503

GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping

APPLICANT: CHONG, Pele

APPLICANT: OOMEN, Raymond P.

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED

TITLE OF INVENTION: PROTEASE ACTIVITY

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
ADDRESSEE: Sim & McBurney
STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: MSG 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/615,271

FILING DATE: 20-JUN-1996

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael I

REGISTRATION NUMBER: 24,973

REFERENCE/DOCKET NUMBER: 1038-580

TELECOMMUNICATION INFORMATION:

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 223 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-615-271-13

Sequence 13, Application US/09074660

Patent No. 6020183

GENERAL INFORMATION:

APPLICANT: LOOSMORE, Sheena M.

APPLICANT: YANG, Yan-Ping

APPLICANT: CHONG, Pele

APPLICANT: OOMEN, Raymond P.

APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: Analog of Haemophilus Hin47 Protein with

REDUCED Protease Activity

NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sim & McBurney

STREET: 6th Floor, 330 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: MSG 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

APPLICATION NUMBER: US 08/296,149
 PRIORITY APPLICATION DATA:
 FILING DATE: 26-AUG-1994
 PRIORITY APPLICATION NUMBER: US 08/278,091
 FILING DATE: 21-JUL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-731 MIS:jb
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-074-659-13

Query Match 100.0%; Score 82; DB 3; Length 223;
 best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEFPFMV 16
 ||||| ||||| |||||
 1 VVGTRAAQGEFPFMV 16

RESULT 10
 US-09-074-659-13
 Sequence 13, Application US/09074659
 Patent No. 6025342

GENERAL INFORMATION:
 APPLICANT: LOOMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: THONG, Pele
 APPLICANT: OOMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.
 TITLE OF INVENTION: Analogs of Haemophilus Hin47 Protein with
 NUMBER OF SEQUENCES: 23
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 6th Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: M5G 1R7
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION NUMBER: US/09/106,168
 FILING DATE:
 CLASSIFICATION:
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 08/615,271
 FILING DATE: 20-JUN-1996
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-825
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-09-106-468-13

Query Match 100.0%; Score 82; DB 3; Length 223;
 best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VVGTRAAQGEFPFMV 16
 ||||| |||||
 1 VVGTRAAQGEFPFMV 16

RESULT 12
US-09-106-466A-13
Sequence 13, Application US/09106466A

GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: COMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sim & McBurney
 STREET: 6th Floor, 330 University Avenue
 CITY: Toronto
 STATE: Ontario
 COUNTRY: Canada
 ZIP: MSG 1R7

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/106,466A
 FILING DATE:
 CLASSIFICATION: 514514

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/615,271
 FILING DATE: 20-JUN-1996
 CLASSIFICATION:
 ATTORNEY/AGENT INFORMATION:
 NAME: Stewart, Michael I.
 REGISTRATION NUMBER: 24,973
 REFERENCE/DOCKET NUMBER: 1038-824

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (416) 595-1155
 TELEFAX: (416) 595-1163
 INFORMATION FOR SEQ ID NO: 13:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 223 amino acids
 TYPE: amino acid
 STRANDEDNESS: Single
 TOPOLOGY: linear

US-09-106-467-13

Query Match 100.0%; Score 82; DB 4; Length 223;
 Best Local Similarity 100.0%; Pred. No. 3.1e-06;
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VVGGTRAQGEPPFMV 16
 Db 1 VVGGTRAQGEPPFMV 16

RESULT 14
US-08-553-516-2

Sequence 2, Application US/08553516

PATENT NO. 5693520

GENERAL INFORMATION:
 APPLICANT: Brammer, Sven
 ATTORNEY/AGENT INFORMATION:
 NAME: Hastrup, Sven
 TITLE OF INVENTION: A Recombinant Trypsin-Like Protease

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5693520 No. 5693520disk of No. 5693520th America, Inc.
 STREET: 405 Lexington Avenue
 CITY: New York
 STATE: NY
 COUNTRY: USA
 ZIP: 10174

COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette
 COMPUTER: IBM Compatible
 OPERATING SYSTEM: DOS
 SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/553,516
 FILING DATE: No. 5693520ember 3, 1995

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
 NAME: Gress, Valera A.
 REGISTRATION NUMBER: 35,127

REFERENCE/DOCKET NUMBER: 3962-204-US

TELECOMMUNICATION INFORMATION:

RESULT 13
US-09-106-467-13
Sequence 13, Application US/09106467

PATENT NO. 6153580

GENERAL INFORMATION:
 APPLICANT: LOOSMORE, Sheena M.
 APPLICANT: YANG, Yan-Ping
 APPLICANT: CHONG, Pele
 APPLICANT: COMEN, Raymond P.
 APPLICANT: KLEIN, Michel H.

TITLE OF INVENTION: ANALOG OF HAEMOPHILUS HIN47 WITH REDUCED NUMBER OF SEQUENCES: 23

CORRESPONDENCE ADDRESS:

TELEPHONE: 212-867-0123
 TELEFAX: 212-878-9655
 TELEX:
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 224 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-553-516-2

Query Match 72.0%; Score 59; DB 1; Length 248;
 Best Local Similarity 62.5%; Pred. No. 0.02%;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 VVGGTRAAQGEPFPFV 16
 Db 25 IVGGTSASAGDFFPFIV 40

Search completed: April 11, 2003, 17:58:50
 Job time : 17 secs

LT 15
 US-08-238-1130-2
 Sequence 2, Application US/08238130
 Patent No. 5702334
 GENERAL INFORMATION:
 APPLICANT: Hastrup, Sven
 APPLICANT: Branner, Sven
 APPLICANT: Jorgensen, Birthe R.
 APPLICANT: Christensen, Tove
 APPLICANT: Jorgensen, Birgitte B.
 APPLICANT: Shuster, Jeffrey R.
 APPLICANT: Madden, Mark
 APPLICANT: Moyer, Donna L.
 APPLICANT: Fuglsang, Claus
 TITLE OF INVENTION: PROCESSES FOR PRODUCING AN ENZYME
 NUMBER OF SEQUENCES: 9
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: No. 5702334o No. 5702934disk of No. 5702934th Americaa, Inc.
 STREET: 405 Lexington Avenue, 64th Floor
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10174-6401
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Tape
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/238,130
 FILING DATE: 04-MAY-1994
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: DK 522/93
 FILING DATE: 05-MAY-1993
 ATTORNEY/AGENT INFORMATION:
 NAME: Agris Dr., Cheryl H.
 REGISTRATION NUMBER: 34,086
 REFERENCE/DOCKET NUMBER: 3,965.200-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 212-867-0123
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 248 amino acids
 TYPE: amino acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 FEATURE:
 NAME/KEY: Protein
 LOCATION: 25..248
 FEATURE:



Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 11, 2003, 17:54:35 ; Search time 35 Seconds
(without alignments)
60.915 Million cell updates/sec

Title: US-10-006-223-1
Perfect score: 82
Sequence: 1 VVGTRAAQGEEPFMV 16

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT;*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT;*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT;*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT;*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT;*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT;*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT;*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT;*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT;*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT;*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT;*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT;*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT;*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT;*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT;*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT;*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT;*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT;*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT;*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT;*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT;*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT;*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT;*

RESULT 1
AAB26847
ID AAB26847 standard; peptide; 15 AA.
XX
AC AAB26847;
DT 29-JAN-2001 (first entry)
XX
DB Protease CGW-3 N-terminal peptide sequence.
XX
KW Protease; CGW-3; plasminogen activator; fibrinolysis;
KW arterial thrombosis.
XX
OS Streptomyces sp.
XX
PN CN1260394-A.
XX
PD 19-JUL-2000.
XX
PP 21-OCT-1999; 99CN-0121864.
XX
PR 21-OCT-1999; 99CN-0121864.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	95.1	15	21 AAB26847	Protease CGW-3 N-terminal peptide sequence.
2	67	81.7	20	21 AAB03088	Protease CGW-3 N-terminal peptide sequence.
3	62	75.6	20	21 AAB03085	Protease CGW-3 N-terminal peptide sequence.
4	61	74.4	258	23 ABB83101	Protease CGW-3 N-terminal peptide sequence.
5	59	72.0	248	15 AAR66999	Protease CGW-3 N-terminal peptide sequence.
6	59	72.0	248	16 AAR84716	Protease CGW-3 N-terminal peptide sequence.
7	59	72.0	248	17 AAR88470	Protease CGW-3 N-terminal peptide sequence.
8	59	72.0	248	21 AAB03659	Protease CGW-3 N-terminal peptide sequence.
9	56	68.3	360	22 ABB60104	Protease CGW-3 N-terminal peptide sequence.
10	54	65.9	251	22 ABB64578	Protease CGW-3 N-terminal peptide sequence.

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Drosophila melanog
L. bimaculatus Plasmi
Guinea Pig AST pro
Epithelin protein. M
Murine epithelin. M
Murine type II mem
Mouse epithelin-like
Crayfish hydrolase
Crayfish protease
N-terminal of a cr
N-terminal of cray
Crayfish protease
Human serine prote
Human membrane-type
Human matrix metalloprote
Drosophila melanog
Porcine AST protei
Human cancer assoc
Human matrixmetalloprote
Human peptidase, H
Human peptidase, H
Human membrane-type
Human protein sequ
Tumour antigen der
Human matrixmetallopro
Human membrane-type
Human peptidase, H
Human TAG-15. Ho
Human membrane-type
Novel human diagno
P. monodon hydrola
Tiger prawn trypti
N-terminal of a sh
Human matrixmetallopro
Panaculus monodon tr
Drosophila melanog
Rabbit AST protein

ALIGNMENTS

Result 1	ID	Peptide sequence
AAB26847	ID AAB26847 standard; peptide; 15 AA.	XX AC AAB26847; DT 29-JAN-2001 (first entry) XX DB Protease CGW-3 N-terminal peptide sequence. XX KW Protease; CGW-3; plasminogen activator; fibrinolysis; KW arterial thrombosis. XX OS Streptomyces sp. XX PN CN1260394-A. XX PD 19-JUL-2000. XX PP 21-OCT-1999; 99CN-0121864. XX PR 21-OCT-1999; 99CN-0121864.

The present invention relates to a protease termed CGW-3 which is obtained by from soil streptomyces strain C3667. CGW-3 is a serine protease and possesses fibrinolysis activity and can activate plasminogen. The N-terminal portion of CGW-3 is represented by the present sequence. The CGW-3 protease can be used in the treatment of arterial thrombosis.

XX SQ Sequence 15 AA;

Query Match 95.1%; Score 78; DB 21; Length 15;
Best Local Similarity 100.0%; Pred. No. 1.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

1 VVGCTRAAAGEFPFM 15
1 VVGCTRAAAGEFPFM 15

```

XX SQ Sequence 15 AA;

RESULT 2
AAB03088 standard; peptide; 20 AA.

AC AAB03088;

XX DT 10-OCT-2000 (first entry)

XX DE N-terminus of Streptomyces griseus trypsin.

XX DE Trypsin; bacterial; serine protease; Trichoderma; coagulant;

XX DE hypertensive; antiinflammatory; leather preparation; silk treatment.

XX DS Streptomyces griseus.

XX PN JP2000116377-A.

XX PD 25-APR-2000.

XX PR 08-OCT-1998;

FT Region /label= Loop-XI
FT Region /label= Loop-XI
FT Region /label= Loop-XI
FT Region /label= Loop-XII
XX WO9530743-A1.
XX 16-NOV-1995.
XX 04-MAY-1995; 95WO-DK00180.
XX 04-MAY-1995; 94DK-0000509.
XX (NOVO) NOVO-NORDISK AS.
PI Fredholm H, Von Der Osten C;
XX WPI: 1995-404110/51.
DR AAT05182.

Variants of a parent trypsin-like Fusarium protease - used as constituents in detergent compositions

FT Claim 1: Page 57-58; 80pp; English.
XX The trypsin-like protease (ARR84716) from Fusarium oxyphorum DSM 2672 includes 13 surface loop regions which can be substituted, esp. by corresponding loop regions II and IV of cattle trypsin 2pin, or loops IV or VI of rat trypsin 1trm, or by other modified loop structures. The protease variants are obt. by site-directed mutagenesis of encoding cDNA (AAT05182). They show improved proteolytic activity and decreased susceptibility to oxidation than the parent enzyme, and are useful in detergent compositions and additives.

XX Sequence 248 AA;

Query Match 72.0%; Score 59; DB 16; Length 248;
Best Local Similarity 62.5%; Pred. No. 0.055; Mismatches 4; Indels 0; Gaps 0;

CC 1 VYGGTAAQGFPEFVY 16
CC :|||:|||:|||:|||:
CC 25 IVGGTSASAGDFPEFIV 40

RESULT 8
ID AAB03659 standard; Protein; 248 AA.
XX AAB03659;
AC AAB03659;
XX 25-SEP-2000 (first entry)
DE Pre-pro-trypsin amino acid sequence.
XX Non-toxic; non-pathogenic; recombinant protein production; protease; pre-pro-trypsin.
XX Fusarium oxyphorum.
XX US56060305-A.
XX 09-MAY-2000.
PD 13-MAR-1997; 97US-0816915.
XX PF 30-JUN-1994; 94US-0269449.
XX PR 15-MAR-1995; 95US-044678.
PR 04-OCT-1996; 96US-0726105.
XX (NOVO) NOVO NORDISK BIOTECH INC.
XX Wendy YT, Shuster JR, Moyer DT, Royer JC;
XX WPI: 2000-349678/30.
DR N-PSDB; AAS53321.

XX New non-pathogenic recombinant fusarium host cell, useful for expressing heterologous proteins especially fungal enzymes such as alkaline endoglucanase or alkaline protease -

XX Disclosure; Column 15-18; 32PP; English.

XX PF 15-JUN-1995; 95WO-US07743.
XX PR 15-MAR-1995; 95US-0404678.

XX The invention relates to a non-toxic, non-pathogenic recombinant Fusarium host cell of the section Discolor, with ATCC accession number 20334. The

CC cell is used in the recombinant production of proteins. The present
 CC sequence represents the *F. oxysporum* pre-pro-trypsin protein sequence,
 CC which is an example of an enzyme which can be expressed by the cells of
 CC the invention. The cells are useful for expressing heterologous proteins
 CC especially fungal enzymes such as alkaline endoglucanase or alkaline
 CC proteases, e.g. *F. oxysporum* pre-pro-trypsin gene, and also hormones,
 CC growth factors and receptors. The cells are non-toxic and are efficient
 CC in the recombinant production of fungal enzymes.

XX SQ Sequence 248 AA;

Query Match 72.0%; Score 59; DB 21; Length 248;
 Best Local Similarity 62.5%; Pred. No. 0.055; Matches 4; Mismatches 2; Indels 0; Gaps 0;
 Matches 10; Conservative 4; Mismatches 2; Indels 2; Gaps 0;

Qy 1 VVGGTRAAQGEFPFMV 16
 :|||: :|||: :|||:
 Db 25 IVGGTSASAGDFPFIV 40

XX LT 9

ID ABB60104 standard; Protein; 360 AA.
 XX WO20011042-A2.
 AC ABB60104;
 XX 26-MAR-2002 (first entry)

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.
 AC 27-SEP-2001.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

DB Drosophila melanogaster polypeptide SEQ ID NO 7104.
 XX WO20011042-A2.

Qy 1 VVGGTRAAQGEFPFMV 15
 :|||: :|||: :|||:
 Db 102 IVGGTRASGKBFPPFM 116

RESULT 1.0
 ABB64578
 ID ABB64578 standard; Protein; 251 AA.
 XX
 AC ABB64578;
 XX DT 26-MAR-2002 (first entry)

DB Drosophila melanogaster polypeptide SEQ ID NO 20526.
 XX KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX OS Drosophila melanogaster.
 XX PN WO20011042-A2.
 XX PD 27-SEP-2001.
 XX PF 23-MAR-2001; 2001IWO-US09231.
 XX PR 23-MAR-2000; 2000US-191637P.
 XX PR 11-JUL-2000; 2000US-0614150.
 XX PA (PEKE) PE CORP NY.
 XX PI Venter JC, Adams M, Li PWD, Myers EW;
 XX DR 2001-656860/75.
 XX DR N-PSDB; ABL08681.

DB New isolated nucleic acid detection reagent for detecting 1000 or more
 CC genes from Drosophila and for elucidating cell signalling and cell-cell
 CC interactions -
 XX Disclosure; SEQ ID NO 20526; 21pp + Sequence Listing; English.

DB The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL0511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB7737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 251 AA;

Query Match 65.9%; Score 54; DB 22; Length 251;
 Best Local Similarity 50.0%; Pred. No. 0.39; Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Qy 1 VVGGTRAAQGEFPFMV 16
 :|||: :|||: :|||:
 Db 28 IVGGTRASGKBFPPFM 43

RESULT 1.1
 ABB61105
 ID ABB61105 standard; Protein; 294 AA.
 XX AC ABB61105;

FT Protein /label= propeptide
 FT 187..418
 FT /label= trypsin-like protein
 XX
 WO200218562-A1.
 PN
 PD 07-MAR-2002.
 XX
 PP 28-AUG-2001; 2001WO-JP07349.
 XX
 PR 28-AUG-2000; 2000JP-0257104.
 PR 05-MAR-2001; 2001JP-0059753.
 PA (TEIJIN LTD.
 XX
 PI Eguchi H, Chokki M, Yamamura S, Mita R, Matsugi T;
 XX
 WPI; 2002-315539/35.
 DR N-PSDB; ABL50743.
 DR
 Airway-specific trypsin-like enzymes for use in diagnosis and screening
 compounds or polypeptides as inhibitors of AST activity, PAR activation
 and mucus production, and judging therapeutic efficacy -
 Example 10; Page 146-148; 165pp; Japanese.
 PS
 The present invention describes an airway-specific trypsin-like enzyme
 (AST) is a protein comprising the whole or a part of the amino acid
 sequence of (ABB05964 (I)), in which a propeptide moiety containing all
 or part of an amino acid sequence of AST between position 1 and
 Arg at position 186 is bonded to a trypsin-like moiety containing Ile at
 position 187 to Ile at position 418 in a 533 amino acid sequence via a
 disulfide bond. The AST enzyme can be used in the diagnosis and screening
 of compounds and polypeptides as inhibitors of AST activity, protease
 activated receptor (PAR) activation, mucus production promotion, cell
 proliferation, calcium flow into cells or EGFR pathway activation by AST
 and judging therapeutic efficacy. The present sequence represents guinea
 pig AST from the present invention.
 XX
 SQ Sequence 418 AA;
 Query Match 65.9%; Score 54; DB 23; Length 418;
 Best Local Similarity 62.5%; Pred. No. 0..9;
 Matches 10; Conservative 4; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 VVGCTRQHGPFPFMV 16
 187 VVGCTQADQDNEPWQV 202
 RESULT 14
 ID AAE23083 standard; Protein; 855 AA.
 XX
 AAE23083;
 DT 21-AUG-2002 (first entry)
 XX
 Epithin protein.
 DE
 PN WO200203787-A2.
 PD 17-JAN-2002.
 XX
 PR 06-JUL-2001; 2001WO-US21422.
 XX
 PR 06-JUL-2000; 2000US-216109P.
 PR 06-JUL-2000; 2000US-216251P.

PR 06-JUL-2000; 2000US-216258P.
 PR 06-JUL-2000; 2000US-216768P.
 PR 10-JUL-2000; 2000US-217449P.
 PR 10-JUL-2000; 2000US-217450P.
 PR 27-JUL-2000; 2000US-217660P.
 PR 27-JUL-2000; 2000US-221491P.
 PR 27-JUL-2000; 2000US-221669P.
 PR 27-JUL-2000; 2000US-221670P.
 PR 07-AUG-2000; 2000US-223170P.
 PR 07-AUG-2000; 2000US-223172P.
 PR 26-OCT-2000; 2000US-223460P.
 PR 26-OCT-2000; 2000US-244037P.
 PR 26-JUN-2001; 2001US-301217P.
 XX
 PA (DELT-) DELTAGEN INC.
 XX
 PI Allen KD, Leviten MW;
 XX
 WPI; 2002-154853/20.
 DR N-PSDB; AAD7039.
 PT Novel non-human transgenic animal, preferably transgenic mice
 comprising disruption in target gene, e.g., trypsin gene, useful for
 identifying an agent that modulates expression or function of target
 gene -
 XX
 Disclosure; Fig 6; 74pp; English.
 XX
 The present invention relates to non-human transgenic animals preferably
 transgenic mice comprising disruption in target gene such as trypsin,
 gene. The invention also relates to compositions and methods relating
 to the characterisation of gene functions. The transgenic animals are
 useful for identifying an agent that modulates the expression or function
 of a target. They are useful for identifying an agent that modulates a
 phenotype associated with a disruption in trypsin genes or limulus
 clotting factor protease-like genes by administering an agent to the
 transgenic animal and determining whether the agent modulates the
 phenotype where the agent has effect on decreased body weight, decreased
 thymus weight, decreased thymus to body weight ratio, increased pre-pulse
 inhibition, significant decrease in their response latency to the hot
 plate test or a decreased response threshold to metrazol. Agents that
 modulate the expression, function or activity of the target gene are
 useful for treating a disorder associated with a mutation in trypsin
 gene or in limulus clotting factor protease-like gene. The transgenic
 animals are useful for testing the efficacy of proposed genetic and
 pharmacological therapies for human genetic diseases. They are useful
 as models for diseases, disorders or conditions associated with
 phenotypes relating to a disruption in a target and to identify drugs,
 pharmaceuticals, therapies and interventions which may be effective in
 treating a disease or other phenotypic characteristics of the animal.
 The present sequence is epithin protein. This sequence is used in the
 exemplification of the invention.

SQ Sequence 855 AA;
 Query Match 64.6%; Score 53; DB 23; Length 855;
 Best Local Similarity 62.5%; Pred. No. 2.2;
 Matches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
 Qy 1 VVGCTRQHGPFPFMV 16
 615 VVGCTQADQDNEPWQV 630
 Db 615 VVGCTQADQDNEPWQV 630

RESULT 15
 AAB98507
 ID AAB98507 standard; Protein; 902 AA.
 XX
 AC AAB98507;
 XX
 DT 03-AUG-2001 (first entry)

DE Murine epithin.
 XX
 KW Murine; TADG-15; cytostatic; vaccine; ovarian tumour; cancer; epithin;
 KW tumour antigen-derived gene 15; serine protease.
 XX
 OS Mus musculus.
 XX
 PN WO200129056-A1.
 XX
 PD 26-APR-2001.
 XX
 PF 20-OCT-2000; 2000WO-US29095.
 XX
 20-OCT-1999; 99US-0421213.

FA (UTAR-) UNIV ARKANSAS.
 XX
 PI O'Brien TJ, Tanimoto H;
 XX
 WPI: 2001-381031/40.

Novel extracellular serine protease, termed tumor antigen-derived gene 15 protein overexpressed in carcinomas and DNA encoding it, for
 PT diagnosis, treatment, prevention of cancer, particularly breast,
 PT ovarian cancer
 PT
 Disclosure: Page 97-99; 130pp; English.

XX The present invention relates to human tumour antigen-derived gene 15
 CC (TADG-15) protein and coding sequence (see AAH22601 and AB98500).
 CC TADG-15 is an extracellular serine protease. It was found that TADG-15 is
 CC over-expressed in ovarian tumours. TADG-15 protein or its fragments of
 CC 9-20 residues that lack TADG-15 protease activity are useful for
 CC vaccinating an individual against TADG-15, having, suspected of having or
 CC at risk of getting cancer. Furthermore, the TADG-15 gene can be used as a
 CC diagnostic or therapeutic target in cancer. The present sequence was used
 CC in a sequence homology alignment with TADG-15.
 XX Sequence 902 AA:

Query Match 64.6%; Score 53; DB 22; Length 902;
 st Local Similarity 62.5%; Pred. No. 2.4;
 tches 10; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 VVGCTRRAOGEFPFMV 16
 Db 615 VVGTRNADEGEWPWQV 630

Search completed: April 11, 2003, 17:57:13
 Job time : 37 secs